

Universidade Federal do Rio Grande do Sul
Instituto de Matemática

Local models and hidden nonlocality in Quantum Theory

Leonardo Guerini de Souza

*Master's thesis written under the supervision
of Alexandre Tavares Baraviera, presented to
Programa de Pós-Graduação em Matemática -
UFRGS as a partial requirement for the title of
Master in Mathematics.*

Porto Alegre, March of 2014.

Master's thesis submitted by Leonardo Guerini de Souza¹² to Programa de Pós-Graduação em Matemática of Universidade Federal do Rio Grande do Sul, as a partial requirement for the title of Master in Mathematics.

Date: April 10, 2014.

Supervisor:

Dr. Alexandre Tavares Baraviera

Examiners:

Dr. Carlos Felipe Lardizabal Rodrigues (IM-UFRGS)

Dr. Marcelo de Oliveira Terra Cunha (DMAT-UFMG)

Dra. Sandra Denise Prado (IF-UFRGS)

¹Supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

²E-mail adress: guerini.leonardo@gmail.com

Abstract

This Master's thesis has two central subjects: the simulation of correlations generated by local measurements on entangled quantum states by local hidden-variables models and the revelation of hidden nonlocality. We present and detail the Werner's local model and the hidden nonlocality of some Werner states of dimension $d \geq 5$, the Gisin-Degorre's local model for a Werner state of dimension $d = 2$ and the local model of Hirsch *et al.* for mixtures of the singlet state and noise, all of them for projective measurements. Finally, we introduce the local model for POVMs of Hirsch *et al.* for a state constructed upon the singlet with noise, that still violates the CHSH inequality after local filters are applied, hence presenting the so-called genuine hidden nonlocality.

Contents

Introduction	3
1 Preliminaries	5
1.1 Dirac's notation	5
1.2 Basic definitions and results	6
1.2.1 Linear algebra	6
1.2.2 Basic Probability Theory	10
2 Quantum Theory	11
2.1 States	11
2.2 Measurements	14
2.2.1 Projective measurements	15
2.2.2 POVMs	17
2.3 The Bloch sphere	17
2.4 The partial trace	18
2.5 EPR	20
3 Nonlocality	22
3.1 The CHSH inequality	22
3.2 The Horodecki Criterion	25
3.3 Local hidden variables	27
3.4 Local models	28
4 Werner's local model	32
4.1 Werner states	32
4.2 Bob's response function and the hidden variables	38
4.3 Alice's response function and a local model for an entangled state	42
4.4 Barrett's local model	45
5 Hidden nonlocality	49

6	Gisin-Degorre's local model	53
6.1	Shared randomness	54
6.2	A local model for the EPR experiment correlations	55
6.3	The sampling	59
6.4	The communication	59
6.5	A local model for the Werner state	60
7	Genuine hidden nonlocality	62
7.1	A local model for dichotomic projective measurements	63
7.2	Revealing nonlocality	66
7.3	A local model for POVMs	66
7.4	Revealing genuine nonlocality	69
7.5	Genuine and maximal	71
	Conclusion	72

Introduction

Quantum Theory is an intrinsically probabilistic theory, that is, when we perform a measurement on a quantum system, we can only tell the probabilities associated to each possible outcome, and not the outcome itself. In the situation where the system is composed by two distinct parties, we can perform a local measurement on each party, generating a distribution of joint probabilities, referent to each outcome obtained. Generally, joint probabilities (originated by quantum measurements or not) are not independent, i.e., cannot be expressed by the product of the probability of each party. However, considering the existence of an extra information, which we call local hidden variables, in certain cases such factorization becomes possible. When we are restricted to correlations generated by measurements on quantum states, the previous sentence can be rewritten as the following: some quantum states admit a local hidden variables model, or, in short, a local model. Those are said to be local states. The construction of a local model is a hard task, and even 25 years after the first of them, created by Reinhard Werner [1], the number of states for which were constructed a local model is still small.

The common characteristic shared by correlations that do not admit a local factorization even when local hidden variables are considered is called nonlocality. Another way to characterize this property is through Bell inequalities [2]. A Bell inequality is a relation that should be satisfied by all local correlations, particularly by all correlations generated by measurements on local states. While the absence of a local model for a given state does not imply its nonlocality, the violation of a single Bell inequality is enough to attest it.

Separable quantum states are naturally local. Maybe just as interesting as the use of local hidden variables to simulate correlations of entangled quantum states is the fact that some local states do violate a Bell inequality when subjected to a sequence of local measurements [3]. In other words, even if there is a local model for a given state, maybe there are local measurements such that one of the possible after-measurement states violates a Bell inequality. Such intermediate measurements are called filters, and a state that allows this possibility is said to present hidden nonlocality. Therefore, the action of the filters is to reveal the state's nonlocality.

The main objective of this dissertation is to present some examples of local models

and some cases of hidden nonlocality. In Chapters 1 and 2 are presented definitions and basic results from the topics around the subject, as well as a short overview of Quantum Theory. This is the starting point to discuss the EPR experiment [4], which motivated the appearance of Bell inequalities.

In Chapter 3 we present the CHSH inequality [5], which is the only Bell inequality employed during the thesis, as well as the Horodecki Criterion, which gives us the largest CHSH violation that a given state provides [6]. We also formally define the role of the local hidden variables [7] and the construction of a local model.

Chapter 4 is turned to Werner’s local model for projective measurements [1], the first local model to arise, which influenced basically all the subsequent works, being one of the most important papers in the area³. We motivate the study of Werner states, which are the ones used for its construction, and present the hidden variables and response functions to be used in each system.

In Chapter 5 we present the work of Sandu Popescu [3], showing that some of the states for which we constructed the Werner’s local model present hidden nonlocality: applying a simple choice of local filters, we are able to obtain a state that violates maximally the CHSH inequality, when the local dimension is $d \geq 5$.

Chapter 6 brings the Gisin-Degorre’s local model [8, 9], capable of simulating the EPR experiment if we allow classical communication between the parties. However, even without communication or any other additional resource, we are able to simulate a Werner state of local dimension $d = 2$, for projective measurements.

Finally, Chapter 7 is based on the work of Flavien Hirsch, Marco Túlio Quintino, Joseph Bowles and Nicolas Brunner [10]. We start by showing the existence of hidden nonlocality in a state of local dimension $d = 2$ and then present an example of the so-called genuine hidden nonlocality: we provide a state which has a local model for POVMs and, nevertheless, violates the CHSH inequality after the application of local filters.

³It was in this paper where first appears the definition of entanglement for mixed states, for example, property until then defined only for pure states.

Chapter 1

Preliminaries

In this chapter we present the standard notation while dealing with Quantum Theory, introduced by Paul Dirac, and a miscellaneous list of basic results and definitions that will be used throughout the thesis.

1.1 Dirac's notation

Dirac's notation is a mnemonic notation that is very useful in the handling with Quantum Theory. The elements of a vector space V are denoted by $|v\rangle$ and the elements of the dual V^* are denoted by $\langle v|$. The symbols $|\cdot\rangle$ and $\langle\cdot|$ already defines the object; the letters only serve as labels. Consequently, the canonical inner product of vectors u and v is simply written as $\langle u|v\rangle \equiv \langle u|v\rangle$.

Along the thesis, sometimes we will want to speak in the adjoint operator T^\dagger of a given operator T (the proper definitions will be given in the next section). While in the traditional notation we write

$$\langle v, Tu\rangle = \langle T^\dagger v, u\rangle,$$

in Dirac's notation both sides of the above expression are written the same,

$$\langle v|(T|u\rangle) = (\langle v|T)|u\rangle = \langle v|T|u\rangle.$$

This happens because the dual element of $T^\dagger|v\rangle$ is exactly $\langle v|T$. Sometimes, to emphasizing this step, we will make use of the abuse of notation

$$\langle v|Tu\rangle := \langle v|T|u\rangle =: \langle T^\dagger v|u\rangle.$$

1.2 Basic definitions and results

1.2.1 Linear algebra

Definition 1.1. A Hilbert space is a pair $(V, \langle \cdot, \cdot \rangle)$, where V is a vector space and $\langle \cdot, \cdot \rangle$ is an inner product that induces a distance function for which V is complete, that is, every Cauchy sequence is convergent.

Except if mentioned otherwise, the inner product that we will be considering is

$$\langle |v\rangle, |w\rangle \rangle = \langle v | w \rangle.$$

Definition 1.2. Let V, W be vector spaces over the same field. The tensor product of them, denoted by $V \otimes W$, is the space generated by vectors of the form $|v\rangle \otimes |w\rangle$, where $|v\rangle \in V$ and $|w\rangle \in W$, which obey to the following relations:

1. $(\lambda |v\rangle) \otimes |w\rangle = |v\rangle \otimes (\lambda |w\rangle) = \lambda(|v\rangle \otimes |w\rangle);$
2. $(|v\rangle + |v'\rangle) \otimes |w\rangle = |v\rangle \otimes |w\rangle + |v'\rangle \otimes |w\rangle;$
3. $|v\rangle \otimes (|w\rangle + |w'\rangle) = |v\rangle \otimes |w\rangle + |v\rangle \otimes |w'\rangle;$
4. If $(V, \langle \cdot, \cdot \rangle_V)$ and $(W, \langle \cdot, \cdot \rangle_W)$ are Hilbert spaces, then $(V \otimes W, \langle \cdot, \cdot \rangle)$ is a Hilbert space, where

$$\langle (|v\rangle \otimes |w\rangle), (|v'\rangle \otimes |w'\rangle) \rangle_{V \otimes W} = \langle v | v' \rangle \langle w | w' \rangle,$$

with $|v'\rangle \in V, |w'\rangle \in W$.

Notation 1. We will use the simplified notation $|v\rangle \otimes |w\rangle \equiv |v\rangle |w\rangle \equiv |vw\rangle$ for vectors of a tensor product of two spaces.

Definition 1.3. Let $A \in \mathcal{L}(\mathcal{H}) = \{T : \mathcal{H} \rightarrow \mathcal{H}; T \text{ is linear}\}$, where \mathcal{H} is a finite dimensional Hilbert space, and $[A] = (a_{ij})$ be a matrix representation of A . The trace of A is the sum of the elements of the main diagonal of $[A]$, that is,

$$\text{tr}(A) = \sum_i a_{ii}.$$

Proposition 1.2.1. Given a d -dimensional Hilbert space \mathcal{H} , let $A, B \in \mathcal{L}(\mathcal{H})$ and $V : \mathcal{H} \otimes \mathcal{H} \rightarrow \mathcal{H} \otimes \mathcal{H}$ be the “flip” linear operator, defined by $V(|ab\rangle) = |ba\rangle$. Then $\text{tr}(VA \otimes B) = \text{tr}(AB)$.

Proof. Fix an orthonormal base $\mathcal{B} = \{|1\rangle, \dots, |d\rangle\}$ for \mathcal{H} . Note that we can write the

matrix representation $A = (a_{ij})_{i,j=1}^d$ as

$$A = \sum_{i,j=1}^d a_{ij} |i\rangle \langle j|.$$

Indeed, $|i\rangle \langle j|$ is the linear operator that vanishes for every vector of \mathcal{B} except $|j\rangle$, hence all the entries of its matrix representation are null except for the one in the j -th column and i -th line. Similarly for B ,

$$B = \sum_{k,l=1}^d b_{kl} |k\rangle \langle l|.$$

Therefore,

$$\begin{aligned} \text{tr}(AB) &= \text{tr} \left(\left[\sum_{i,j=1}^d a_{ij} |i\rangle \langle j| \right] \left[\sum_{k,l=1}^d b_{kl} |k\rangle \langle l| \right] \right) \\ &= \text{tr} \left(\sum_{i,j,l}^d a_{ij} b_{jl} |i\rangle \langle l| \right) \\ &= \sum_{i,j,l}^d a_{ij} b_{jl} \text{tr}(|i\rangle \langle l|) \\ &= \sum_{i,j}^d a_{ij} b_{ji}. \end{aligned}$$

In the same line, the tensor product becomes

$$A \otimes B = \sum_{i,j,k,l} a_{ij} b_{kl} |ki\rangle \langle lj|. \quad (1.1)$$

As for the flip operator, since it can be seen as a permutation over the elements of the base of $\mathcal{H} \otimes \mathcal{H}$, we can describe the action of its matrix representation on another matrix as a simply permutation of the matrix lines. We can express this fact as

$$V \cdot |ij\rangle \langle kl| = (V |ij\rangle) \langle kl| = |ji\rangle \langle kl|. \quad (1.2)$$

Now we are ready to compute $\text{tr}(VA \otimes B)$. Using Eqs. (1.1) and (1.2), we have

$$\begin{aligned}\text{tr}(VA \otimes B) &= \text{tr} \left(\sum_{i,j,k,l} a_{ij} b_{kl} V(|ki\rangle \langle lj|) \right) \\ &= \sum_{i,j,k,l} a_{ij} b_{kl} \text{tr}(|ik\rangle \langle lj|) \\ &= \sum_{i,j} a_{ij} b_{ji},\end{aligned}$$

thus equal to $\text{tr}(AB)$. \square

Notation 2. We will denote a $n \times n$ diagonal matrix $A = (a_{ij})$ (i.e., a matrix A for which $a_{ij} = 0$ for all $i \neq j$) by $A = \text{diag}(a_{11}, \dots, a_{nn})$.

Proposition 1.2.2. Given a Hilbert space \mathcal{H} , an unit vector $|\psi\rangle \in \mathcal{H}$ and a linear operator A acting on \mathcal{H} , we have

$$\text{tr}(A|\psi\rangle \langle \psi|) = \langle \psi| A |\psi\rangle.$$

Proof. Consider the orthogonal basis $\{|e_i\rangle\}$ of \mathcal{H} , where $|e_1\rangle = |\psi\rangle$, and the matrix representation $[A]$ of A in this basis. The first column of $[A]$ is the vector $A|\psi\rangle = A|e_1\rangle = \sum_i a_i |e_i\rangle$. Thus

$$\langle \psi| A |\psi\rangle = \langle e_1| A |e_1\rangle = \langle e_1| \left(\sum_i a_i |e_i\rangle \right) = a_1 \langle e_1| e_1\rangle = a_1.$$

On the other hand, the matrix representation of the projector $|\psi\rangle \langle \psi|$ in this basis is $\text{diag}(1, 0, \dots)$, so the product $A|\psi\rangle \langle \psi|$ equals to the matrix which the first column is the vector $A|\psi\rangle$ and the rest of matrix elements equals to zero, resulting that $\text{tr}(A|\psi\rangle \langle \psi|) = a_1$. \square

Definition 1.4. An operator $T \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is said to be a product operator if $T \in \mathcal{L}(\mathcal{H}_A) \otimes \mathcal{L}(\mathcal{H}_B) = \{T_A \otimes T_B \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B); T_A \in \mathcal{L}(\mathcal{H}_A), T_B \in \mathcal{L}(\mathcal{H}_B)\}$.

Proposition 1.2.3. Every non-product operator can be written as a linear combination of product operators.

Proof. The proposition follows from the fact that if $\{|a_i\rangle\}$ is a basis for \mathcal{H}_A and $\{|b_j\rangle\}$ is a basis for \mathcal{H}_B , then $\{|a_i b_j\rangle \langle a_{i'} b_{j'}|\} = \{|a_i\rangle \langle a_{i'}| \otimes |b_j\rangle \langle b_{j'}|\}$ is a basis for $\mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$. \square

Since $\mathcal{L}(\mathcal{H})$ is also a vector space, the tensor product $\mathcal{L}(\mathcal{H}_A) \otimes \mathcal{L}(\mathcal{H}_B)$ is well defined. We can define the action of $T_A \otimes T_B : \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B) \rightarrow \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$ by

$$T_A \otimes T_B(M_A \otimes M_B) = (T_A M_A) \otimes (T_B M_B)$$

for product operators and extend to non-product operators by linearity.

Definition 1.5. Let $A \in \mathcal{L}(\mathcal{H})$ and $[A]$ the matrix representation of A with respect to the base \mathcal{B} . The adjoint of A is the operator $A^\dagger \in \mathcal{L}(\mathcal{H})$ whose matrix representation with respect to \mathcal{B} is $[A]$ transposed and complex conjugated.

Definition 1.6. Let $T \in \mathcal{L}(\mathcal{H})$. T is said to be

- Hermitian if $T = T^\dagger$;
- normal if $TT^\dagger = T^\dagger T$;
- positive semi-definite if $\langle \psi | T | \psi \rangle \geq 0$, $\forall |\psi\rangle \in \mathcal{H}$;
- unitary if $T^\dagger = T^{-1}$.

The next two basic results about these classes of operators we will only enounce; one can find the proofs in [11].

Theorem 1. (i) Any positive semi-definite operator on a vector space is a Hermitian operator.

(ii) Any Hermitian operator on a vector space is a normal operator.

(iii) Any unitary operator on a vector space is a normal operator.

Theorem 2. [Spectral Decomposition Theorem] An operator on a vector space is normal if and only if is diagonal with respect to some basis for the space.

Proposition 1.2.4. Let $T, P \in \mathcal{L}(\mathcal{H})$ such that $P = |v\rangle \langle v|$ is a projector. Then $PT = TP$ if and only if $|v\rangle$ is an eigenvector of T .

Proof. Suppose that $TP = PT$. Then

$$T|v\rangle = TP|v\rangle = PT|v\rangle = \lambda|v\rangle,$$

for some $\lambda \in \mathbb{C}$, by definition of P .

Conversely, suppose $T|v\rangle = \lambda|v\rangle$ and take $|u\rangle \in \mathcal{H}$. Consider the base $\{|v_i\rangle\}$ of \mathcal{H} such that $|v\rangle = |v_1\rangle$. Then $|u\rangle = \sum_i a_i |v_i\rangle$, with $a_i \in \mathbb{C}$, and

$$TP|u\rangle = T(a_1|v\rangle) = a_1T|v\rangle = a_1\lambda|v\rangle.$$

On the other hand,

$$PT|u\rangle = P(a_1\lambda|v\rangle + \sum_{i>1} a_i T|v_i\rangle) = a_1\lambda|v\rangle,$$

completing the proof. □

Corollary 1. Let $S, T \in \mathcal{L}(\mathcal{H})$ such that S is Hermitian and its spectral decomposition is $\sum_i a_i P_i$. If $TP_i = P_i T$, $\forall i$, then $T = \sum_i b_i P_i$.

1.2.2 Basic Probability Theory

Definition 1.7. Let A, B be random variables. The conditional probability that $B = b$ given that $A = a$ is defined by

$$p(B = b|A = a) = \frac{p(A = a, B = b)}{p(A = a)}.$$

When $p(A = a) = 0$ we make the convention that $p(B = b|A = a) = 0$.

Definition 1.8. Random variables A, B are said to be independent if $p(A = a, B = b) = p(A = a)p(B = b)$.

Notation 3. We often denote $p(A = a, B = b)$ by $p(a, b)$, leaving the “ $A =$ ” and “ $B =$ ” implicit.

Theorem 3 (Law of total probability). If A, B are random variables, then

$$p(b) = \sum_a p(b|a) p(a),$$

where the sum is over all values a that A can assume.

Definition 1.9. The expectation or expected value of a random variable A that take values in \mathbb{R} is defined by

$$E(A) = \sum_a p(a)a$$

where the sum is over all values a that A can assume.

Proposition 1.2.5. The expectation has the following properties.

- (i) $E(A)$ is linear in A .
- (ii) If A, B are independent, then $E(AB) = E(A)E(B)$.

We recommend [12] for further definitions and results.

Chapter 2

Quantum Theory

This chapter has no intention to be a didactic introduction to Quantum Theory and should not be the first text about the subject to be readed. Its purpose is to present the minimum of the mathematical framework related to the quantum operations and phenomena that we are interested in this text, such as measurements and nonlocality. Therefore, no physical motivation will be exposed and various of basic and important topics will be completely ignored (such as time evolution of quantum systems, for example). For this reason, the postulates of Quantum Theory will be exposed as definitions. On the other hand, some very specific results will have to appear.

We will start directly making use of the density operator formalism. For an introduction to Quantum Theory and its formalisms we suggest references [13], [14] and [15].

2.1 States

In Quantum Theory, we postulate that a system is associated to a Hilbert space \mathcal{H} .¹ Considering the set $\mathcal{L}(\mathcal{H})$ of linear operators of \mathcal{H} , a state is an element of $\mathcal{L}(\mathcal{H})$ that describes completely the system.

Definition 2.1. *A state of a system associated to \mathcal{H} is an operator $\rho \in \mathcal{L}(\mathcal{H})$ which is positive semi-definite with unit trace.*

The subset of $\mathcal{L}(\mathcal{H})$ formed by states is denoted by $\mathcal{D}(\mathcal{H})$, where the “ \mathcal{D} ” comes from the expression “density operator”, another term for “quantum state” in this formalism. $\mathcal{D}(\mathcal{H})$ is a convex set, that is, every convex combination of density matrices is also a density matrix. If the state ρ is a one-dimensional projector (that is, if $\rho^2 = \rho$ and $\text{Im}(\rho)$ has dimension 1), then $\rho = |\psi\rangle\langle\psi|$ for some $|\psi\rangle \in \mathcal{H}$, thus we can identify the density operator ρ with the vector $|\psi\rangle$. The unit trace condition implies that $|\psi\rangle$ has euclidean

¹Since throughout this text the Hilbert spaces will have finite dimension, we can think that each system we are going to speak about is associated to \mathbb{C}^d , for some dimension d .

norm equals to 1. Every other kind of density operator is called a *mixed state* and can be written as a convex combination of projectors, i.e.,

$$\rho \in \mathcal{D}(\mathcal{H}) \implies \rho = \sum_i p_i |\psi_i\rangle \langle \psi_i|, \quad (2.1)$$

with unit $|\psi_i\rangle \in \mathcal{H}$ and $p_i \geq 0$ satisfying $\sum_i p_i = 1$. Notice that a one-dimensional projector is a mixed state with only one term on the sum. Therefore, it is called a *pure state*. However, the decomposition in (2.1) is not unique. Take for example

Example 1.

$$\rho = \frac{3}{4} |0\rangle \langle 0| + \frac{1}{4} |1\rangle \langle 1| = \frac{1}{2} |a\rangle \langle a| + \frac{1}{2} |b\rangle \langle b|,$$

where $|a\rangle = \frac{\sqrt{3}}{2} |0\rangle + \frac{1}{2} |1\rangle$ and $|b\rangle = \frac{\sqrt{3}}{2} |0\rangle - \frac{1}{2} |1\rangle$.

In this thesis, we will turn our attention mostly to composite systems; generally, we will be considering two different systems \mathcal{H}_A and \mathcal{H}_B . The right way to describe such compositions is through the tensor product of Hilbert spaces, which is a Hilbert space itself.

Definition 2.2. *The state space of a composite system is the tensor product of the state spaces of the component systems.*

In the product $\mathcal{H}_A \otimes \mathcal{H}_B$, the first factor will be said to be Alice's system and the second, Bob's system, following the usual terminology used in the literature.

In a composite system, we can speak about separability.

Definition 2.3. *A pure state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ is said to be a product state if there are $|\psi_A\rangle \in \mathcal{H}_A$ and $|\psi_B\rangle \in \mathcal{H}_B$ such that*

$$|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle. \quad (2.2)$$

The pure states that are not product are said to be entangled.

The concept of product state is just the restriction of the term product operator to the set $\mathcal{D}(\mathcal{H})$. However, the analogy does not hold for Proposition 1.2.3, since it is false that every state is a convex combination of product states.

As any definition given by a denial, it is usually hard to decide whether a state is entangled or not: in principle, one needs to show that equation (2.2) does not hold for any pair $|\psi_A\rangle \otimes |\psi_B\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$.

Example 2. *Let $\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^2$ and $\mathcal{B} = \{|0\rangle, |1\rangle\}$ be a basis of \mathbb{C}^2 . The state*

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |01\rangle)$$

of \mathcal{H} is a product state, since

$$|\psi\rangle = |0\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle).$$

Example 3. Let \mathcal{H} and \mathcal{B} be the same as above but take

$$|\Phi_+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle).$$

If $|\Phi_+\rangle$ were separable, then there would be $|a\rangle, |b\rangle \in \mathbb{C}^2$ such that

$$|\Phi_+\rangle = |a\rangle \otimes |b\rangle. \quad (2.3)$$

Writing these vectors in terms of \mathcal{B} , we have

$$|a\rangle = \alpha |0\rangle + \beta |1\rangle, \quad |b\rangle = \gamma |0\rangle + \delta |1\rangle,$$

for some $\alpha, \beta, \gamma, \delta \in \mathbb{C}^2$. Thus

$$|a\rangle \otimes |b\rangle = \alpha\gamma |00\rangle + \alpha\delta |01\rangle + \beta\gamma |10\rangle + \beta\delta |11\rangle.$$

According to equation (2.3), this implies

$$\begin{aligned} \alpha\gamma &= \beta\delta = \frac{1}{\sqrt{2}} \\ \alpha\delta &= \beta\gamma = 0, \end{aligned}$$

which is an insoluble system. Therefore, we conclude that $|\Phi_+\rangle$ is entangled. (In fact, $|\Phi_+\rangle$ is one of the so-called Bell states, which are maximally entangled states of $\mathbb{C}^2 \otimes \mathbb{C}^2$.)

Generalizing the above concepts, we have the following definition.

Definition 2.4. A state ρ of a composite system $\mathcal{H}_A \otimes \mathcal{H}_B$ is said to be separable if

$$\rho = \sum_i p_i \rho_{A,i} \otimes \rho_{B,i}$$

for some $\rho_{A,i} \in \mathcal{H}_A$ and $\rho_{B,i} \in \mathcal{H}_B$, with $\sum_i p_i = 1$ and $p_i \geq 0$. A state which is not separable is said to be entangled.

2.2 Measurements

We saw how to represent systems and states. Now we will define how an observer can interact with the systems, that is, how a measurement takes place.

Definition 2.5. *Given a state $\rho \in \mathcal{H}$, a measurement on ρ is a set of operators $\{M_i\}$ that acts on \mathcal{H} and satisfy the completeness relation*

$$\sum_i M_i^\dagger M_i = I.$$

The index i refers to the outcomes that may occur in the measurement; outcome i occurs with probability

$$\text{tr}(M_i \rho M_i^\dagger)$$

and the state of the system after the measurement is

$$\frac{M_i \rho M_i^\dagger}{\text{tr}(M_i \rho M_i^\dagger)}.$$

It is easy to see that the completeness relation implies that the probabilities of the outcomes indeed sum to one.

The outcomes of a measurement can be attached to each measurement operator in an arbitrary way; in general, they do not have a special meaning. For projective measurements however, we will see that this is usually done in a somewhat natural way.

The fact that we use the trace to calculate the probabilities is known as Born's rule. In order to formalize and to strengthen the axiomatic basis of Quantum Theory, Andrew Gleason showed that every probability measure on the set of closed subspaces of a Hilbert space with dimension greater than two is given by the trace of its projector times a positive semi-definite operator with trace one [16]. In some sense, this result legitimates the use of the Born's rule.

The next definition can be found in [14].

Definition 2.6. *Let \mathcal{V} be a vector space and \mathcal{C} the set of closed subspaces of \mathcal{V} . A probability measure on \mathcal{C} is a measure $\mu : \mathcal{C} \rightarrow [0, 1]$ such that $\mu(\mathcal{V}) = 1$ and if $\{E_i\} \subset \mathcal{C}$ is a countable collection of mutually orthogonal subspaces that generates $E \subset \mathcal{C}$ then*

$$\mu(E) = \sum_i \mu(E_i)$$

Theorem 4 (Gleason). *Let \mathcal{V} be a vector space over \mathbb{C} with dimension $d > 2$ and \mathcal{C} the set of closed subspaces of \mathcal{V} . If μ is a probability measure on \mathcal{C} , then there is a positive*

semi-definite operator $\rho \in \mathcal{L}(\mathcal{V})$ with unit trace such that

$$\mu(E) = \text{tr}(P\rho),$$

for all $E \in \mathcal{C}$, where P is the projector onto E .

The proof of Gleason's theorem can be found in his original work [16] and in a simplified version in [17]. In [18], the author discusses other theorems that formalize Born's rule.

We will see later that Gleason's theorem can give us a hint for the construction of local models.

2.2.1 Projective measurements

A simple and important kind of measurements occurs when the measurement operators $\{P_i\}$ are projectors, i.e., $P_i^2 = P_i$, and satisfies $P_i P_j = \delta_{ij} P_i$. In this case, the completeness relation becomes

$$\sum_i P_i = I$$

and the probabilities are given by

$$\text{tr}(P_i \rho P_i^\dagger) = \text{tr}(P_i \rho P_i) = \text{tr}(P_i^2 \rho) = \text{tr}(P_i \rho).$$

If the state is pure, then

$$\text{tr}(P_i \rho) = \text{tr}(P_i |\psi\rangle \langle \psi|) = \langle \psi | P_i | \psi \rangle,$$

where the last equality is given by Proposition 1.2.2.

While a general measurement can be associated with a simple partition of the identity I , in the projective case the association can go further, identifying each projective measurement $\{P_i\}$ with the Hermitian operator

$$M = \sum_i m_i P_i$$

(here in its spectral decomposition form), acting on the space state of the system. In this case, the measurement operators are the projectors onto its eigenvectors and the possible outcomes are its eigenvalues. Such Hermitian operator M is called an *observable*. The set of observables of a system \mathcal{H} will be denoted by $\mathcal{O}(\mathcal{H}) \subset \mathcal{L}(\mathcal{H})$.

Example 4. Important examples of observables are the Pauli matrices,

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Together with the identity $I_{2 \times 2}$, the Pauli matrices span the real vector space $H_2(\mathbb{C})$ of the 2×2 Hermitian matrices.

The measurement of observable σ_i is referred as “measurement of spin along the i axis”, for $i = x, y, z$.

Example 5. More generally, given $|v\rangle = (v_x, v_y, v_z)$ a real three-dimensional unit vector, we can define the observable²

$$v \cdot \sigma \equiv v_x \sigma_x + v_y \sigma_y + v_z \sigma_z = \begin{pmatrix} v_z & v_x - i v_y \\ v_x + i v_y & -v_z \end{pmatrix},$$

which has eigenvalues ± 1 . Hence, a projective measurement related to $v \cdot \sigma$ has possible outcomes ± 1 and projectors onto the corresponding eigenspaces are given by

$$P_{\pm} = \frac{I \pm v \cdot \sigma}{2}$$

and thus the corresponding probabilities of the measurement applied to the pure state $|\psi\rangle$ are

$$\begin{aligned} p(\pm 1) &= \text{tr} \left(|\psi\rangle \langle \psi| \frac{I \pm v \cdot \sigma}{2} \right) \\ &= \frac{1 \pm \langle \psi | v \cdot \sigma | \psi \rangle}{2}. \end{aligned}$$

The measurement of this observable is referred to as “measurement of spin along the $|v\rangle$ direction”.

Example 6. Take the pure state

$$|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$

and let’s measure the observable σ_z . As said above, its eigenvalues are ± 1 associated to eigenvectors $|0\rangle, |1\rangle$, with respective projectors (measurement operators) $P_{\pm}^z = (I \pm \sigma_z)/2$. Hence, the possible resulting states post-measurement are $|0\rangle$ and $|1\rangle$, occurring with probability $\text{tr}(P_-^z |+\rangle \langle +|) = 1/2 = \text{tr}(P_+^z |+\rangle \langle +|)$.

²We denote $|a\rangle \cdot |b\rangle$ simply by $a \cdot b$, in order to do not overload the notation.

On the other side, $|+\rangle$ is itself an eigenvector of the observable σ_x , together with $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$. Notice that we can also choose the opposite way and write the “ z -basis” $\{|0\rangle, |1\rangle\}$ in terms of the “ x -basis” $\{|+\rangle, |-\rangle\}$:

$$|0\rangle = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle), \quad |1\rangle = \frac{1}{\sqrt{2}}(|+\rangle - |-\rangle).$$

Therefore, measuring observable σ_x on any of the states $|0\rangle, |1\rangle$, the probability to obtain any of the possible resulting states $|-\rangle, |+\rangle$ is $\text{tr}(P_-^x |+\rangle \langle +|) = 1/2 = \text{tr}(P_+^x |+\rangle \langle +|)$, where the measurement operators are $P_\pm^x = (I \pm \sigma_x)/2$.

We conclude from these observations that measuring observable σ_z on an eigenvector of σ_x leads to a resulting state uniformly distributed on $\{|0\rangle, |1\rangle\}$; but measuring σ_x on any of those states leads to a resulting state uniformly distributed on $\{|+\rangle, |-\rangle\}$. Thus it is impossible for a particle to have spin determined simultaneously in both directions x and z . Though it perhaps seems an unimportant conclusion, this fact will play a decisive role in Section 2.5.

2.2.2 POVMs

Observe that, since the probabilities of a measurement $\{M_i\}$ are given by $\text{tr}(M_i \rho M_i^\dagger) = \text{tr}(\rho M_i^\dagger M_i)$, we can associate to the measurement the positive operators $\{Q_i\}$, where $Q_i = M_i^\dagger M_i$. Although such operators are not enough to determine the resulting post-measurement state, many times we are only interested in the probabilities $\text{tr}(\rho Q_i)$, and we can restrain our attentions to the simplified description of the measurement provided by the $\{Q_i\}$. These operators are known as the *POVM elements* associated to the measurement (sometimes called *effects*), where the acronym holds for positive operator-value measure.

2.3 The Bloch sphere

The simplest quantum system that we can imagine is the one associated to the Hilbert space \mathbb{C}^2 , in which the pure states are unit vectors of the form $|v\rangle = \alpha|0\rangle + \beta|1\rangle$ and are called *qubits*. As we imply, the vectors $|0\rangle, |1\rangle$ form an orthogonal base for \mathbb{C}^2 .

Because $|v\rangle$ has norm 1, we must have $|\alpha|^2 + |\beta|^2 = 1$. Writing $\alpha = \alpha_1 + i\alpha_2$ and $\beta = \beta_1 + i\beta_2$, this condition becomes $\alpha_1^2 + \alpha_2^2 + \beta_1^2 + \beta_2^2 = 1$, so in principle we see that the set of qubits is isomorphic to the sphere $\mathbb{S}^3 \subset \mathbb{R}^4$.

However, given an observable of $\mathcal{L}(\mathbb{C}^2)$, its eigenvectors are orthogonal, thus they form a base for \mathbb{C}^2 . The coefficients of a pure state written in such base are related to the probabilities with which the resulting post-measurement state becomes the corresponding

eigenvector.

Example 7. If $|v\rangle = \alpha|0\rangle + \beta|1\rangle$ and we measure σ_z (whose eigenvectors are $|0\rangle$ and $|1\rangle$), then

$$\text{tr}(|0\rangle\langle 0| |v\rangle\langle v|) = |\alpha|^2, \quad \text{tr}(|1\rangle\langle 1| |v\rangle\langle v|) = |\beta|^2,$$

that is, the resulting state is $|0\rangle$ with probability $|\alpha|^2$ and $|1\rangle$ with probability $|\beta|^2$.

Hence, since a state describes the system by giving the probabilities of the possible outcomes of any measurement, we can identify the pure states $|v\rangle$ and $e^{i\theta}|v\rangle$: the $e^{i\theta}$ factor will not influence in the computation of probabilities $|\alpha|^2$ and $|\beta|^2$. Another way to say this is that $|v\rangle$ and $e^{i\theta}|v\rangle$ are physically indistinguishable.

This identification tells us that we have no need of four real coefficients to describe the set of pure states physically distinct, only three will suffice³. Thus, the set of physically distinct qubits is isomorphic to \mathbb{S}^2 .

As we saw in Section 2.2.1, to each real three-dimensional unit vector $|v\rangle$ can be associated the observable $v \cdot \sigma$. Now we know that each of such vectors can also be associated uniquely to a pure state. In order to distribute the pure states $|v\rangle$ in \mathbb{S}^2 , we can take each of them to correspond to the direction of the eigenvector associated to $+1$ in the measurement of the observable $v \cdot \sigma$. The sphere \mathbb{S}^2 together with this correspondence is called *the Bloch sphere*.

Notice that the pair of vectors that are orthogonal in \mathbb{R}^3 are collinear in the Bloch sphere: the states corresponding to vectors $|0\rangle, |1\rangle$ points to the north and south poles, and $|-\rangle, |+\rangle$ are in opposite points of the equator of the Bloch sphere. We can think that, in the process of allocation of states corresponding to vectors from \mathbb{R}^3 in the Bloch sphere, the angles between the vector and the positive y -axis get doubled. This mean that if we consider a scalar product over the Bloch sphere, it will be given by

$$|a\rangle \cdot |b\rangle = \cos(2\theta),$$

where θ is the angle between the vectors in \mathbb{S}^2 .

2.4 The partial trace

Definition 2.7. Let $\mathcal{H}_A \otimes \mathcal{H}_B$ be a composite Hilbert space. We define the partial trace (in relation to \mathcal{H}_B) $\text{tr}_B : \mathcal{L}(\mathcal{H}_A) \otimes \mathcal{L}(\mathcal{H}_B) \rightarrow \mathcal{L}(\mathcal{H}_A)$ by

$$\text{tr}_B(M_A \otimes M_B) = M_A \cdot \text{tr}(M_B)$$

³The right way to formalize this claim is by showing that the referred identification determines a equivalence relation and considering the quotient space generated by it [14].

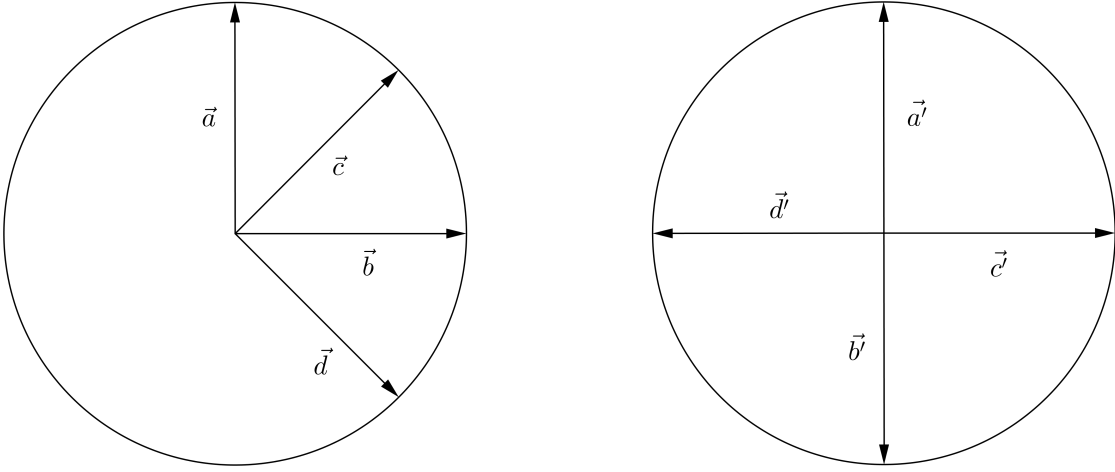


Figure 2.1: Suppose the vectors \vec{a}, \vec{b} are $|0\rangle, |1\rangle$ and vectors \vec{c}, \vec{d} are $|+\rangle, |-\rangle$. In the left side we have them represented in a section of the sphere \mathbb{S}^2 and in the right, in a section of the Bloch sphere.

for product operators and extend to non-product operators by linearity. Analogously, we define the partial trace in relation to subsystem \mathcal{H}_A .

Given a state of a composite system $\mathcal{H}_A \otimes \mathcal{H}_B$, we can find the state that describe each subsystem via partial trace. Such state is called *reduced density operator*.

Definition 2.8. Let ρ be the state that describes the composite system $\mathcal{H}_A \otimes \mathcal{H}_B$. Then the reduced density operator

$$\rho_A = \text{tr}_B(\rho)$$

describes subsystem \mathcal{H}_A .

The description provided by the partial trace referred above is about measurements: suppose that Alice shares the state ρ with Bob and perform a measurement $\{M_i\}$. Then $\text{tr}_B(\rho) = \rho_A$ is the only state that satisfies

$$\text{tr}(M_a \otimes I \rho) = \text{tr}(M_a \rho_A)$$

for any a and $\{M_i\}$, that is, is the only state that provides the right probability of obtain outcome a when we consider a measurement only on Alice's system. In other words, concerning to local measurements only on Alice's side, to say that Alice shares state ρ with Bob is the same as to say that Alice holds the state $\text{tr}_B(\rho)$.

Example 8. Suppose that Alice and Bob shares the singlet state $|\Psi_-\rangle \langle \Psi_-|$. Then Alice's reduced state is

$$\begin{aligned}
\text{tr}_B(|\Psi_-\rangle \langle \Psi_-|) &= \frac{1}{2} \text{tr}_B(|01\rangle \langle 01| + |10\rangle \langle 10| - |01\rangle \langle 10| - |10\rangle \langle 01|) \\
&= \frac{1}{2}(|0\rangle \langle 0| \text{tr}(|1\rangle \langle 1|) + |1\rangle \langle 1| \text{tr}(|0\rangle \langle 0|) - |0\rangle \langle 1| \text{tr}(|1\rangle \langle 0|) \\
&\quad - |1\rangle \langle 0| \text{tr}(|0\rangle \langle 1|)) \\
&= \frac{1}{2}(|0\rangle \langle 0| + |1\rangle \langle 1|) \\
&= \frac{I_{2 \times 2}}{2}.
\end{aligned}$$

2.5 EPR

In objection to the view where measurements in quantum systems do not properly reveal some already existing information, but rather just lead the system to probabilistically assume one of the possible resulting states, Albert Einstein, together with Nathan Rosen and Boris Podolsky proposed a thought experiment [4]. Their idea was to show, using entangled states, that the description of reality provided by Quantum Theory was not complete. Here, we are going to use the simplified formulation given by David Bohm [19].

The first step is to understand what the authors propose as description of reality.

Definition 2.9. *An element of reality is a physical property that can be determined precisely (i.e., with probability equal to unity) without disturbing the system. A theory is complete if contains a counterpart for every element of reality.*

EPR consider a bipartite system where the parts, Alice and Bob, share the singlet state

$$|\Psi_-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle).$$

The singlet state has the property⁴ that, given any direction $|v\rangle$, it can be written as

$$|\Psi_-\rangle = \frac{1}{\sqrt{2}}(|v_+v_-\rangle - |v_-v_+\rangle),$$

where $|v_+\rangle, |v_-\rangle$ are eigenstates of the observable $v \cdot \sigma$, related to eigenvalues ± 1 . This means that the outcomes of any measurement of $v \cdot \sigma$ in both parts will be perfectly anti-correlated: if Alice obtains +1, then Bob obtains -1, and vice-versa.

Now imagine that Alice and Bob are far away from each other, Alice measures σ_z and obtains outcome +1. So, if Bob also measures σ_z in his particle, we know that the

⁴This property is derived from the fact that the singlet has total spin equal to zero, which roughly means that each part must have its spin pointing to exactly opposite directions [19].

resulting outcome will be -1. Alice's measurement cannot instantaneously disturb Bob's system because of the distance separating them, but still we can predict with probability 1 the value of σ_z . Thus, according to Definition 2.9, σ_z is an element of reality of Bob's system.

We can imagine that, instead of σ_z , Alice measures σ_x and using the same argument conclude that σ_x is an element of reality of Bob's system as well. (Does not really matter if Alice indeed measured σ_z or σ_x or did not do any measurement at all: the important fact is that it would be possible to predict the property 'spin along the z -axis' or 'spin along the x -axis' of Bob's particle.) Therefore, there exist elements of reality corresponding to those properties simultaneously. A complete theory should, therefore, simultaneously assign values for such properties.

However, we saw in Example 6 that Quantum Theory forbids both observables to be determined at the same time⁵. The conclusion of EPR was that Quantum Theory is not complete. Despite being a good approximation of experimental data, they believed that should exist another theory, satisfying their completeness criterium without paradoxes.

It is important to emphasize that we have taken two major assumptions to insure this conclusion: the *realism* (the intrinsic existence of values for elements of reality, independent of measurements) and the *locality* (distant measurements cannot influence each other instantaneously). Hence, the EPR argument proves that *if* Quantum Theory were a locally realistic theory, *then* would not be a complete one.

In the next chapter, we will see that Quantum Theory is not locally realistic (or, as we will simply say from now on, *local*).

⁵This rest upon the fact that σ_z and σ_x do not commute. Any other pair of non-commuting observables would do the job, as we will consider in Chapter 6.

Chapter 3

Nonlocality

Nearly thirty years after Ref. [4] was published, John Bell made a proposal that allows one to experimentally test whether Quantum Theory is locally realistic or not [2]. The main tool used for performing such test was a linear inequality for the joint probabilities (or equivalently, the joint expectations) of both parts, the first of a class of inequalities now known as *Bell inequalities*. To obey the Bell inequalities is a condition that a local-realist theory should satisfy. As we will see, Quantum Theory is able to violate such inequalities, consequently annulling the incompleteness' conclusion of EPR.

Though Bell have developed the first Bell inequality in 1964, the most popular one is the CHSH inequality, dating from 1969.

3.1 The CHSH inequality

The CHSH inequality [5], named after its authors, Clauser, Horne, Shimony and Holt, is the only one that we will have to keep in mind for the discussions proposed in this thesis.

Suppose that we have two parts, Alice and Bob, and a referee, a third part which is not related to any system, but is able to prepare two particles and distribute them to Alice and Bob, repeating the procedure an unlimited number of times. Once each part has its particle, it must choose among two possible measurements, say M_Q or M_R for Alice and M_S or M_T for Bob. They do not know in advance which measurement they will perform; by receiving the particle, each of the two decides it randomly. Also, each measurement M_i has outcome i , for $i = Q, R, S, T$, that takes values on $\{1, -1\}$.

Making the assumption of realism, we will consider that Q (and similarly for R, S, T) is an objective property of Alice's particle, being merely revealed by the measurement M_Q . The course of action is arranged in such a way that Alice and Bob do their measurements at the precise same time. Thus, assuming locality, Alice's measurement cannot disturb the result of Bob's measurement, and vice-versa.

We will now consider the quantity

$$QS + RS + RT - QT.$$

Since the above expression can be rewritten as

$$Q(S - T) + R(S + T)$$

and $S, T \in \{1, -1\}$, we must have $Q(S - T) = 0$ or $R(S + T) = 0$. In either case, we have

$$QS + RS + RT - QT \in \{+2, -2\}.$$

Now, if $p(q, r, s, t)$ is the probability that right before the measurements are performed the system is in a state that $Q = q, R = r, S = s$ and $T = t$, then the expectation becomes

$$\begin{aligned} E(QS + RS + RT - QT) &= \sum_{q,r,s,t \in \{+1, -1\}} p(q, r, s, t)(qs + rs + rt - qt) \\ &\leq 2 \sum_{q,r,s,t \in \{+1, -1\}} p(q, r, s, t) \\ &= 2. \end{aligned}$$

On the other hand, expectation is linear,

$$E(QS + RS + RT - QT) = E(QS) + E(RS) + E(RT) - E(QT).$$

Combining the previous equations, we obtain the CHSH inequality [5]

$$E(QS) + E(RS) + E(RT) - E(QT) \leq 2. \quad (3.1)$$

With this inequality in hands, Alice and Bob are able to compare whether or not the expected values predicted by their theory obeys the inequality. Depending on the result of the comparison (more specifically, if the inequality is violated), they conclude that their theory is nonlocal. They can also proceed to perform a series of measurements in the above conditions on a shared state. The outcomes obtained by each part can then be put together, and the integrated data can be used to approximate each of the expected values in question. If they get an inequality violation, they will be experimentally confirming the nonlocality of Nature itself. If they check that the data obtained match their predictions, they will be showing that Nature agrees with their nonlocal theory.

We now proceed to prove that Quantum Theory is not local.

Example 9. Suppose that Alice and Bob share the singlet state

$$|\Psi_{-}\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}.$$

Consider now the observables

$$M_Q = \sigma_z, \quad M_R = \sigma_x$$

for Alice and

$$M_S = \frac{-\sigma_z - \sigma_x}{\sqrt{2}}, \quad M_T = \frac{\sigma_z - \sigma_x}{\sqrt{2}}$$

for Bob.

The observables M_Q, M_S have eigenvalues $+1, -1$, with respective eigenvectors $|0\rangle, |1\rangle$ for M_Q and $|0\rangle + (-1 - \sqrt{2})|1\rangle, |0\rangle + (-1 + \sqrt{2})|1\rangle$ for M_S . Their projectors onto the eigenspaces are

$$\begin{aligned} P_+^Q &= |0\rangle\langle 0| \\ P_-^Q &= |1\rangle\langle 1| \\ P_+^S &= \frac{|0\rangle\langle 0| + (3 + 2\sqrt{2})|1\rangle\langle 1| + (-1 - \sqrt{2})[|0\rangle\langle 1| + |1\rangle\langle 0|]}{4 + 2\sqrt{2}} \\ P_-^S &= \frac{|0\rangle\langle 0| + (3 - 2\sqrt{2})|1\rangle\langle 1| + (-1 + \sqrt{2})[|0\rangle\langle 1| + |1\rangle\langle 0|]}{4 - 2\sqrt{2}}. \end{aligned}$$

Hence, Quantum Theory predicts that

$$\begin{aligned} E(QS) &= p(qs = 1) - p(qs = -1) \\ &= [p(q = 1, s = 1) + p(q = -1, s = -1)] - [p(q = +1, s = -1) \\ &\quad + p(q = -1, s = +1)] \\ &= [\text{tr}(P_+^Q \otimes P_+^S |\Psi_{-}\rangle\langle \Psi_{-}|) + \text{tr}(P_-^Q \otimes P_-^S |\Psi_{-}\rangle\langle \Psi_{-}|)] \\ &\quad - [\text{tr}(P_+^Q \otimes P_-^S |\Psi_{-}\rangle\langle \Psi_{-}|) + \text{tr}(P_-^Q \otimes P_+^S |\Psi_{-}\rangle\langle \Psi_{-}|)] \\ &= \left[\frac{3 + 2\sqrt{2}}{4(2 + \sqrt{2})} + \frac{1}{4(2 - \sqrt{2})} \right] - \left[\frac{3 - 2\sqrt{2}}{4(2 - \sqrt{2})} + \frac{1}{4(2 + \sqrt{2})} \right] \\ &= \frac{1}{\sqrt{2}}. \end{aligned}$$

Similarly, we may find that

$$E(RS) = E(RT) = \frac{1}{\sqrt{2}}, \quad E(QT) = -\frac{1}{\sqrt{2}}.$$

Thus,

$$E(QS) + E(RS) + E(RT) - E(QT) = 2\sqrt{2} > 2.$$

Since the CHSH inequality is violated, we conclude that Quantum Theory is nonlocal.

3.2 The Horodecki Criterion

Notice that the CHSH inequality (and similarly, all other Bell inequalities) can be seen as a superoperator, whose entries are a quantum state and four observables. In order to test a violation, even after fixing a state, we may achieve different values for the expectation by varying the set of observables.

In Example 9, we saw that, for the singlet state, there exist observables M_Q, M_R, M_S, M_T for which the CHSH is violated. This, however, is not the case for all quantum states. In 1995, Ryszard Horodecki, Pawel Horodecki and Michal Horodecki presented a necessary and sufficient condition for an arbitrary quantum state of $\mathbb{C}^2 \otimes \mathbb{C}^2$ to violate the CHSH inequality and made explicit the greatest value achieved by that state [6].

Associated to the CHSH inequality, there is the observable

$$\mathcal{B}_{CHSH}(x, x', y, y') = x \cdot \sigma \otimes (y + y') \cdot \sigma + x' \cdot \sigma \otimes (y - y') \cdot \sigma,$$

called the CHSH *Bell operator*, where $|x\rangle, |x'\rangle, |y\rangle, |y'\rangle$ are unit vectors in \mathbb{R}^3 . The inequality in question thus becomes

$$E_\rho(\mathcal{B}_{CHSH}) \leq 2,$$

where the subscript emphasize the referred state. Our objective is to check if ρ violates it, hence we want to maximize the expectation over all Bell operators and show that such quantity is achieved by some Bell operator \mathcal{B}_{max} :

$$\sup_{\mathcal{B}_{CHSH}} E_\rho(\mathcal{B}_{CHSH}) = E_\rho(\mathcal{B}_{max}).$$

The Horodecki start by showing that for all $\rho \in \mathcal{D}(\mathbb{C}^2 \otimes \mathbb{C}^2)$, there are vectors $|r\rangle, |s\rangle \in \mathbb{R}^3$ such that

$$\rho = \frac{1}{4} \left(I \otimes I + (r \cdot \sigma) \otimes I + I \otimes (s \cdot \sigma) + \sum_{m,n=1}^3 t_{nm} \sigma_n \otimes \sigma_m \right),$$

where $\sigma_x = \sigma_1, \sigma_y = \sigma_2$ and $\sigma_z = \sigma_3$ are the Pauli matrices and $t_{nm} = \text{tr}(\rho \sigma_n \otimes \sigma_m)$.

Consider the 3×3 matrix T_ρ composed by the coefficients t_{mn} . A simple calculation gives us

$$E_\rho(\mathcal{B}_{CHSH}) = \langle x | T_\rho(|y\rangle + |y'\rangle) + \langle x' | T_\rho(|y\rangle - |y'\rangle).$$

By noticing that the vectors $|y\rangle + |y'\rangle, |y\rangle - |y'\rangle$ are orthogonal and satisfy $\| |y\rangle +$

$|y'\rangle\|^2 + \||y\rangle - |y'\rangle\|^2 = 4$, where $\|\cdot\|$ is the Euclidian norm, we can introduce the pair of unit and mutually orthogonal vectors $|z\rangle, |z'\rangle$ such that

$$|y\rangle + |y'\rangle = 2 \cos \theta |z\rangle, \quad |y\rangle - |y'\rangle = 2 \sin \theta |z'\rangle,$$

for some $\theta \in [0, \pi]$. Then we have

$$\sup_{\mathcal{B}_{CHSH}} E_\rho(\mathcal{B}_{CHSH}) = \sup_{\theta, |x\rangle, |x'\rangle, |z\rangle, |z'\rangle} 2 [\langle x | T_\rho | z \rangle \cos \theta + \langle x' | T_\rho | z' \rangle \sin \theta] \quad (3.2)$$

$$= \sup_{\theta, |z\rangle, |z'\rangle} 2 [\|T_\rho |z\rangle\| \cos \theta + \|T_\rho |z'\rangle\| \sin \theta] \quad (3.3)$$

$$= \sup_{|z\rangle, |z'\rangle} 2 \sqrt{\|T_\rho |z\rangle\|^2 + \|T_\rho |z'\rangle\|^2} \quad (3.4)$$

Now, consider the positive matrix $U_\rho = T_\rho^\dagger T_\rho$ and denote its two greatest eigenvalues by u, \tilde{u} . Using Lagrange multipliers, it is possible to show that

$$\sup_{\{|j\rangle, |k\rangle\}} (\|T_\rho |j\rangle\|^2 + \|T_\rho |k\rangle\|^2) = u + \tilde{u} =: M(\rho),$$

where the supreme is taken over all orthonormal subsets $\{|j\rangle, |k\rangle\} \subset \mathbb{R}^3$. Eq. (3.2) shows us that

$$\sup_{\mathcal{B}_{CHSH}} E_\rho(\mathcal{B}_{CHSH}) = 2\sqrt{M(\rho)}.$$

Conversely, one can take in turn $|z_{max}\rangle, |z'_{max}\rangle$ as the U_ρ eigenvectors maximalizing $M(\rho)$; $|x_{max}\rangle, |x'_{max}\rangle$ as unit vectors in the directions $T_\rho |z_{max}\rangle, T_\rho |z'_{max}\rangle$; and θ_{max} defined by $\|T_\rho |z_{max}\rangle\| \sin \theta_{max} = \|T_\rho |z'_{max}\rangle\| \cos \theta_{max}$. Using these parameters to construct the observable \mathcal{B}_{max} , we have

$$2\sqrt{M(\rho)} = E_\rho(\mathcal{B}_{max}) = \sup_{\mathcal{B}_{CHSH}} E_\rho(\mathcal{B}_{CHSH}).$$

The above reasoning outlines the proof of the following result.

Theorem 5. *There is a CHSH Bell operator \mathcal{B}_{max} such that*

$$E_\rho(\mathcal{B}_{max}) = \sup_{\mathcal{B}_{CHSH}} E_\rho(\mathcal{B}_{CHSH}) = 2\sqrt{M(\rho)}.$$

The straightforward corollary yielded is the announced Horodecki Criterion.

Corollary 2 (Horodecki Criterion). *A state $\rho \in \mathcal{D}(\mathbb{C}^2 \otimes \mathbb{C}^2)$ violates the CHSH inequality if and only if $M(\rho) > 1$.*

3.3 Local hidden variables

The EPR argument can be generalized to a “Bell scenario”, by considering an arbitrary number of systems, possible measurements and possible outcomes. What is central in such kind of experiment is that - let’s suppose the number of systems is two - Alice and Bob both receive from a referee a shared state and give rise to joint probabilities $p(ab|x, y)$, where $a(b)$ is a possible outcome for $x(y)$, one of the possible measurements available to be performed by Alice (Bob). As we saw, in general we have

$$p(a, b|x, y) \neq p(a|x) p(b|y),$$

that is, the outcomes on both sides are not always independent from each other. The existence of such correlations does not necessarily imply a direct influence of one system on the other, though. It can be the case that it is only being revealed a dependence relation between the two systems established in the past, when they may have interacted [7].

If that is the case, there is a set of past factors, described by some variables λ , which influences the outcomes, explaining completely the dependence between them. This means that we must be able to write

$$p(a, b|x, y, \lambda) = p(a|x, \lambda) p(b|y, \lambda), \quad (3.5)$$

standing for the fact that the only needed information to determine the probability of outcome a occurring is the local measurement x in question and the past variables λ ; the distant measurement performed by Bob has nothing to do with it.

Since we are not claiming to have any knowledge over λ besides its existence, in principle it may involve physical quantities that are not completely controllable. Thus, λ will not necessarily be constant for each run of the experiment, even if the referee is careful enough to use the exact same process to prepare the states. Thus, we must consider also a probability distribution $\pi(\lambda)$ over Λ , the set where these variables inhabit, that rules the different values of λ for different experiments.

Eq. (3.5) together with this considerations lead to

$$p(a, b|x, y) = \int_{\Lambda} \pi(\lambda) p(a|x, \lambda) p(b|y, \lambda). \quad (3.6)$$

Another observation that should be made is that the distribution $\pi(\lambda)$ is supposed to be independent of the chosen measurements, i.e., for any choice of x, y , the λ variables are distributed in the same way. In other words, $\pi(\lambda)$ must *a priori* take in consideration all possible measurements on the system.

By using Eq. (3.6), it is possible to rewrite the same expectations of Example 9 in terms of λ , and then derive the CHSH inequality (3.1), as is done in [7]. That is, the assumption of this past variables λ can play the part of the local realism assumption, formulated by EPR. Based on that, we can take Eq. (3.6) as the condition for locality. The variables λ are called *local hidden variables*: ‘local’ because their knowledge enable us to make a local factorization of the joint probabilities, ‘hidden’ because they are apart from the state¹.

3.4 Local models

Strictly speaking, locality is a feature of families of probabilities distributions $\{p(a, b|x, y) \in \mathbb{R} | a \in A, b \in B, x \in X, y \in Y\}$, where A, B are the sets of possible outcomes and X, Y the sets of possible measurements of each part. So, when we say that Quantum Theory is nonlocal, what we really mean is that there exists at least one bipartite state and one pair of measurements that provide a set of correlations which cannot be written in the factorized manner of Eq. (3.6), even when local hidden variables are taken in account. This is guaranteed by a violation of a Bell inequality, like we saw in Example 9.

However, this is not the case for all quantum states. Indeed, for product states the correlations factorize naturally.

Proposition 3.4.1. *Let $A, B \in \mathcal{L}(\mathcal{H})$. Then $\text{tr}(A \otimes B) = \text{tr}(A) \text{tr}(B)$.*

Proof. Since²

$$A \otimes B = \sum_{i,j,k,l} a_{ij} b_{kl} |ki\rangle \langle lj|,$$

we have

$$\begin{aligned} \text{tr}(A \otimes B) &= \sum_{i,j,k,l} a_{ij} b_{kl} \text{tr}(|ki\rangle \langle lj|) \\ &= \sum_{i,k} a_{ii} b_{kk} \\ &= \sum_i a_{ii} \sum_k b_{kk} \\ &= \text{tr}(A) \text{tr}(B). \end{aligned}$$

□

¹We are considering the terms locality and local realism to have basically the same meaning. For a discussion on those terminologies, check [20].

²Check the proof of Proposition 1.2.1.

With Proposition 3.4.1 in mind, we can easily see that a measurement on a product state $\rho^A \otimes \rho^B$ can always be locally factorized:

$$p(a, b | \{M_i\}, \{N_j\}) = \text{tr}(M_a \rho^A \otimes N_b \rho^B) = \text{tr}(M_a \rho^A) \text{tr}(N_b \rho^B) = p(a | \{M_i\}) p(b | \{N_j\}).$$

It is as if no local variables were hidden, or if those who were make no difference in the outcomes: $p(a|x, \lambda) = p(a|x)$ (and the same for Bob). Thus the locality condition is trivially achieved: for any measurements x, y we have

$$\begin{aligned} p(a, b | x, y) &= p(a|x) p(b|y) \\ &= \int_{\Lambda} \pi(\lambda) p(a|x) p(b|y) d\lambda \\ &= \int_{\Lambda} \pi(\lambda) p(a|x, \lambda) p(b|y, \lambda) d\lambda. \end{aligned}$$

On the other side, there is no reason to exclude the existence of an entangled state for which the locality condition is satisfied as well. In principle, it may be that the relation between the systems represented by the entanglement can also be explained by some set of local hidden variables. This would mean that, setting good choices of objects to play the part of the hidden variables $\lambda \in \Lambda$, a probability distribution $\pi(\lambda)$ of these objects and a manner to compute the probabilities of the outcomes generated by each part, we would be able to artificially simulate the correlations $p(ab|x, y)$ in a way that Eq. (3.6) would always be satisfied, no matter which measurements x, y were being considered. In the case where we manage to perform all these tasks successfully, then we say that we have created a *local hidden variables model* for such state.

Definition 3.1. Let $\mathcal{M}(\mathcal{H}) = \{\{M_i\} \subset \mathcal{L}(\mathcal{H}); \sum_i M_i^\dagger M_i = I\}$ be the set of generalized measurements over \mathcal{H} , $O_M \subset \mathbb{R}$ the set of possible outcomes for the measurement $M = \{M_i\}$ and Λ the set of hidden variables. A function

$$\begin{aligned} f^M : O_M \times \Lambda \times \{\{M_i\}\} &\rightarrow [0, 1] \\ (a, \lambda, \{M_i\}) &\mapsto f^M(a, \lambda, \{M_i\}) \end{aligned}$$

is said to be a response function for M if for every fixed $\lambda \in \Lambda$ we have

$$\sum_{j \in O_M} f^M(j, \lambda, \{M_i\}) = 1.$$

A function $f : \mathbb{R} \times \Lambda \times \mathcal{M}(\mathcal{H}) \rightarrow [0, 1]$ is said to be a response function if every restriction

$$\begin{aligned} f|_M : O_M \times \Lambda \times \{\{M_i\}\} &\rightarrow [0, 1] \\ (a, \lambda, \{M_i\}) &\mapsto f(a, \lambda, \{M_i\}) \end{aligned}$$

is a response function for M , that is,

$$f|_M = f^M, \quad \forall M \in \mathcal{M}(\mathcal{H}).$$

The response functions depend on the outcomes, the hidden variables and the whole measurement being applied³, since their task is to attribute a probability for each outcome.

Definition 3.2. A state $\rho \in \mathcal{H}_A \otimes \mathcal{H}_B$ admits a local hidden variables model for projective measurements if there exists a set Λ of hidden variables, a probability distribution $\pi(\lambda)$ and response functions p_A, p_B such that for any pair of observables $P = \sum_i i P_i$, $Q = \sum_j j Q_j$ measured by Alice and Bob, respectively, the equality

$$\text{tr}(\rho P_a \otimes Q_b) = \int_{\Lambda} \pi(\lambda) p_A(a, \lambda, \{P_i\}) p_B(b, \lambda, \{Q_j\}) d\lambda, \quad (3.7)$$

is attained.

Notice that left hand side of the above equality stands for $p(a, b | P_a, Q_b)$, so we have there a twin of Eq. (3.6), where $p_A(a, \lambda, P) = p(a | \lambda, P)$ and $p_B(b, \lambda, Q) = p(b | \lambda, Q)$.

There is no reason for us to restrain the idea of locality to projective measurements. The definition below is a stronger version of Def. 3.2, since projectors are a particular case of positive operators.

Definition 3.3. Analogously to Def. 3.2, a state $\rho \in \mathcal{H}_A \otimes \mathcal{H}_B$ admits a local hidden variables model for POVMs if Eq. (3.7) is satisfied for any pair of POVMs $\{M_i\}, \{N_j\}$, that is,

$$\text{tr}(\rho M_a \otimes N_b) = \int_{\Lambda} \pi(\lambda) p_A(a, \lambda, \{M_i\}) p_B(b, \lambda, \{N_j\}) d\lambda.$$

We will usually use the shortcut expression ‘local model’ to refer to local hidden variables model. In the same fashion, sometimes we will refer to a state which admits a local model as simply ‘local’.

The locality of product states discussed above naturally leads to a local model for a

³Indeed, if the response functions depend only on the specific measurement operator regarding the outcome of interest, Gleason’s theorem would imply that the correlations being reproduced would belong to a separable state. See Section 4.3 for further details.

separable (mixed) state. If our state is

$$\rho = \sum_{i=1}^n p_i \rho_i^A \otimes \rho_i^B$$

then the probabilities we would like to reproduce have the form

$$\text{tr}(M_a \otimes N_b \rho) = \sum_{i=1}^n p_i \text{tr}(M_a \rho_i^A) \text{tr}(N_b \rho_i^B).$$

The comparison of the integral in Eq. (3.7) and the sum in the right side of the above equation practically solve the problem: we just have to take the hidden variables as $\lambda \in \Lambda = \{1, \dots, n\}$ and $\pi(\lambda) = p_\lambda$. The response functions will be same for both parts, the “quantum” response function given by the trace of the state times the measurement operator. Hence, the role of the hidden variables here is only to determine which of the product states present in the mixture we are going to use in the response function.

Since separability implies locality, by contraposition, we conclude

$$\textit{nonlocality} \Rightarrow \textit{entanglement}.$$

In [21], Nicolas Gisin showed that for every pure bipartite entangled state, it is possible to construct observables for which the CHSH inequality is violated. Later, in Ref. [22] and [24] it has been proved that this extends to every multipartite pure state. In other words, for pure states we have the equivalence

$$\textit{nonlocality} \iff \textit{entanglement}.$$

The great and perhaps counterintuitive advance acquired by Reinhard Werner was to show that for mixed states the situation is not the same [1]. That is, it does exist entangled states which admit local models, and thus violate no Bell inequality! So, in general, we have

$$\textit{entanglement} \not\Rightarrow \textit{nonlocality}.$$

That is the first big result that we are going to present, in the next chapter.

Chapter 4

Werner's local model

This whole chapter is devoted to detail Werner's paper [1], in which the first local model appears.

There are two main difficulties in constructing an example of entangled state which admits a local model for projective measurements. The first is to prove that such state is indeed entangled, i.e., it can not be written as a convex combination of separable states. The second is to verify that Eq. (3.7) holds independently of P, Q , which become an infinite system of equations indexed by the set of projective measurements. We shall circumvent both difficulties by considering states of very high symmetry, nominated Werner states.

4.1 Werner states

Definition 4.1. *A state $W \in \mathcal{D}(\mathcal{H} \otimes \mathcal{H})$ is said to be a Werner state if it is $U \otimes U$ -invariant, i.e., $U \otimes UWU^\dagger \otimes U^\dagger = W$, $\forall U \in \mathcal{L}(\mathcal{H})$ unitary.*

The $U \otimes U$ -invariance can be physically interpreted as the property that allows both Alice and Bob to apply the same unitary transformation in their part of the system and still remain with the same global state.

In order to characterize the Werner states, we start asking which operators, not necessarily states, present the $U \otimes U$ -invariance. It is clear, by the definition of unitary, that the identity I has this property. For the *flip* operator $V : \mathcal{H} \otimes \mathcal{H} \rightarrow \mathcal{H} \otimes \mathcal{H}$, defined by

$$V |ij\rangle = |ji\rangle$$

on the product states and extended by linearity, we have $V = V^{-1} = V^\dagger$. It follows that

$$VA \otimes BV^\dagger = B \otimes A$$

and thus

$$\begin{aligned} (U \otimes U)V(U^\dagger \otimes U^\dagger) &= V \Leftrightarrow (U \otimes U)V(U^\dagger \otimes U^\dagger)V^{-1} = I \\ \Leftrightarrow (U \otimes U)(VU^\dagger \otimes U^\dagger V^\dagger) &= I \Leftrightarrow (U \otimes U)(U^\dagger \otimes U^\dagger) = I, \end{aligned}$$

so V is also $U \otimes U$ -invariant. By linearity, we have that any linear combination of these two operators presents this property. The first important result about Werner states is that those are all the operators that have the $U \otimes U$ -invariance.

Proposition 4.1.1. *An operator $A \in \mathcal{L}(\mathbb{C}^d \otimes \mathbb{C}^d)$ is $U \otimes U$ -invariant if and only if A is a linear combination of the identity I and the flip operator V .*

Proof. One of the implications was established above. To the other way, consider A $U \otimes U$ -invariant and the orthonormal basis $\mathcal{B} = \{|e_1e_1\rangle, |e_1e_2\rangle, \dots, |e_de_d\rangle\}$ of $\mathbb{C}^d \otimes \mathbb{C}^d$. The matrix of A associated to this basis has entries $\langle e_je_k | A | e_le_m \rangle$. Using the invariance of A under unitaries U_r which take $|e_r\rangle$ to $-|e_r\rangle$ leaving the other elements fixed, we have

$$\langle e_je_k | A | e_le_m \rangle = \langle e_je_k | U_r^\dagger \otimes U_r^\dagger A U_r \otimes U_r | e_le_m \rangle = \langle U_r \otimes U_r e_je_k | A U_r \otimes U_r | e_le_m \rangle$$

for $r = 1, \dots, d$. Such matrix elements vanishes unless the indices j, k, l, m are equal by pairs. Indeed, if we have $j \notin \{k, l, m\}$, then

$$\langle e_je_k | A | e_le_m \rangle = \langle U_j \otimes U_j e_je_k | A U_j \otimes U_j | e_le_m \rangle = -\langle e_je_k | A | e_le_m \rangle.$$

In the same way, making use of the unitaries \tilde{U}_r that take $|e_r\rangle$ to $i|e_r\rangle$ and leaves the other elements fixed, we see that the terms of the form $\langle e_je_j | A | e_le_l \rangle$ also vanishes.

We conclude that the only non-zero matrix elements have the form $\langle e_je_j | A | e_je_j \rangle$, $\langle e_je_k | A | e_je_k \rangle$ or $\langle e_je_k | A | e_ke_j \rangle$. Since any transposition of two of the basis elements can be realised unitarily, we can apply unitaries $U_{(i,k)}$ that transpose $|e_i\rangle$ and $|e_k\rangle$ and leaves the rest of elements fixed to see that, for any $k \in \{1, \dots, d\}$,

$$\langle e_ie_i | A | e_ie_i \rangle = \langle U_{(i,k)} \otimes U_{(i,k)} e_ie_i | A U_{(i,k)} \otimes U_{(i,k)} | e_ie_i \rangle = \langle e_ke_k | A | e_ke_k \rangle.$$

Hence,

$$\langle e_ie_i | A | e_ie_i \rangle = \gamma, \quad \forall i,$$

for some $\gamma \in \mathbb{C}$. The same unitaries together with $U_{(j,l)}$, defined similarly, gives us

$$\begin{aligned} \langle e_ie_j | A | e_ie_j \rangle &= \alpha, \quad \forall i, j, \quad i \neq j \\ \langle e_ie_j | A | e_je_i \rangle &= \beta, \quad \forall i, j, \quad i \neq j \end{aligned}$$

with $\alpha, \beta \in \mathbb{C}$. The unitaries of the form $U_{(i,k)(j,l)}$ tell us that

$$\langle e_i e_j | A | e_k e_l \rangle = \langle e_k e_l | A | e_i e_j \rangle,$$

and hence all entries are real, that is, actually we have $\alpha, \beta, \gamma \in \mathbb{R}$.

At this point, we already know that A is described by

$$A | e_i e_i \rangle = \gamma | e_i e_i \rangle, \quad A | e_i e_j \rangle = \alpha | e_i e_j \rangle + \beta | e_j e_i \rangle.$$

Using unitaries¹ H_{ij} such that

$$H_{ij} | i \rangle = \frac{| i \rangle + | j \rangle}{\sqrt{2}}, \quad H_{ij} | j \rangle = \frac{| i \rangle - | j \rangle}{\sqrt{2}}$$

and the rest remains fixed, we see that $\gamma = \alpha + \beta$. Indeed, in the particular case where $d = 2$, we have

$$[A] = \begin{pmatrix} \gamma & 0 & 0 & 0 \\ 0 & \alpha & \beta & 0 \\ 0 & \beta & \alpha & 0 \\ 0 & 0 & 0 & \gamma \end{pmatrix}, \quad [H \otimes H] = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix},$$

thus

$$[H \otimes H A H^\dagger \otimes H^\dagger] = \begin{pmatrix} \frac{\gamma+\alpha+\beta}{2} & 0 & 0 & \frac{\gamma-\alpha-\beta}{2} \\ 0 & \frac{\gamma+\alpha-\beta}{2} & \frac{\gamma-\alpha+\beta}{2} & 0 \\ 0 & \frac{\gamma-\alpha+\beta}{2} & \frac{\gamma+\alpha-\beta}{2} & 0 \\ \frac{\gamma-\alpha-\beta}{2} & 0 & 0 & \frac{\gamma+\alpha+\beta}{2} \end{pmatrix},$$

which equals $[A]$ if and only if $\gamma = \alpha + \beta$.

Now, we only have to notice that the non-zero matrix elements of the identity have the form $| e_j e_k \rangle \langle e_j e_k |$, for $j, k = 1, \dots, d$ and the non-zero matrix elements of the flip are $| e_j e_k \rangle \langle e_k e_j |$, with $j, k = 1, \dots, d$. So, when we sum a multiple of the identity and a multiple of the flip, the only entries that overlap are $\langle e_j e_j | e_j e_j \rangle$.

Since this describes all non-zero matrix elements of A , we conclude that $A = \alpha I + \beta V$. \square

This result is valid for all operators $U \otimes U$ -invariants, in particular for Werner states. So, for each of these states, there are parameters α, β such that

$$W = \alpha I + \beta V.$$

¹This is a variation of the Hadamard operator $H = H_{01}$.

With the constraint that a state has unitary trace and using $\text{tr}(I) = d^2$ and $\text{tr}(V) = d$ (the flip operator permutes the basis elements, standing fixed only $|e_i e_i\rangle$, $i = 1, \dots, d$), we are able to bond both parameters together,

$$1 = \text{tr}(W) = \alpha \text{tr}(I) + \beta \text{tr}(V) = \alpha d^2 + \beta d,$$

and thus

$$\alpha = \frac{(1 - \beta d)}{d^2}. \quad (4.1)$$

So we need only one parameter to determine a Werner state. In order to study which Werner states are entangled, we will make use of the parameter ϕ defined by

$$\begin{aligned} \phi &= \text{tr}(WV) = \text{tr}([\alpha I + \beta V]V) = \alpha \text{tr}(V) + \beta \text{tr}(V^2) \\ &= \alpha d + \beta d^2 = \frac{(1 - \beta d)}{d} + \beta d^2 = \frac{1 - \beta d + \beta d^3}{d}, \end{aligned}$$

where we have used that $V^2 = I$ and Eq. (4.1). This way we obtain

$$\beta = \frac{d\phi - 1}{d^3 - d}, \quad \alpha = \frac{d - \phi}{d^3 - d}$$

and therefore

$$W = \frac{(d - \phi)I + (d\phi - 1)V}{d^3 - d}. \quad (4.2)$$

Thus the task of constructing a local model for Werner states is the task of showing that the integral in (3.7) is equal to

$$\text{tr}(WP_a \otimes Q_b) = \frac{(d - \phi) \text{tr}(P_a) \text{tr}(Q_b) + (d\phi - 1) \text{tr}(P_a Q_b)}{d^3 - d}, \quad (4.3)$$

where we have used the formulas $\text{tr}(R \otimes S) = \text{tr}(R) \text{tr}(S)$ and $\text{tr}(VA \otimes B) = \text{tr}(AB)$ (Propositions 1.2.1, 3.4.1).

The next results show how useful this parametrization is to study the entanglement of a Werner state.

Lemma 1. *If $\rho \in \mathcal{D}(\mathcal{H})$, then $\text{tr}(V\rho) \in [-1, 1]$, where V is the flip operator.*

Proof. By noticing that $V^\dagger = V$, we see that all eigenvalues λ of V are real. Since $V^2 = I$, we have

$$V|v\rangle = \lambda|v\rangle \implies |v\rangle = V^2|v\rangle = \lambda^2|v\rangle \implies \lambda^2 = 1,$$

hence $\lambda \in \{\pm 1\}$.

Therefore, the least value achieved by $\text{tr}(V\rho)$ is $\text{tr}(-\rho) = -1$, corresponding to the

situation where $\rho = |\psi\rangle\langle\psi|$ and $|\psi\rangle$ is an eigenvector of V associated with -1 . Similarly, the largest value of $\text{tr}(V\rho)$ is 1, obtained when $\rho = |\psi\rangle\langle\psi|$ and $|\psi\rangle$ is an eigenvector associated with 1. \square

Theorem 6. *The flip operator V is the optimal entanglement witness for Werner states. That is, if $W \in \mathcal{D}(\mathbb{C}^d \otimes \mathbb{C}^d)$ is a Werner state, then it is separable if and only if $\phi = \text{tr}(VW) \geq 0$.*

Proof. Note that

$$\text{tr}(V(|ab\rangle\langle ab|)) = \text{tr}((V|ab\rangle)\langle ab|) = \text{tr}(|ba\rangle\langle ab|) = \delta_{ab},$$

i.e., equals 1 if $|a\rangle = |b\rangle$ and 0 otherwise. Since a separable state is a convex combination of product states of the $|ab\rangle\langle ab|$ form, we conclude that

$$\phi = \text{tr}(VA) \in [0, 1]$$

whenever A is a separable state. This proves the first part of the theorem.

In order to prove the converse, recall that the Werner states are uniquely determined by the parameter ϕ , which lies in $[-1, 1]$ according to Lemma 1. Firstly, we observe that it is enough to show that the Werner states W_0, W_1 , corresponding to $\phi = 0$ and $\phi = 1$, are separable. Indeed, assuming this, then for every $\phi \in [0, 1]$ the corresponding Werner state W_ϕ is given by

$$W_\phi = (1 - \phi)W_0 + \phi W_1.$$

This follows from the linearity of the $U \otimes U$ conjugation (which implies the $U \otimes U$ -invariance of W_ϕ), the achievement of the $\text{tr}(W_\phi) = 1$ condition and from the calculation

$$\text{tr}(VW_\phi) = (1 - \phi)\text{tr}(VW_0) + \phi\text{tr}(VW_1) = \phi.$$

Thus we conclude that W_ϕ is separable, since it is a convex combination of separable states.

The separability of W_0 and W_1 will be proved by using the map

$$\mathbb{P} : A \mapsto \int (U \otimes U)A(U^\dagger \otimes U^\dagger)dU,$$

where dU denotes the Haar measure of the unitary transformations group of \mathbb{C}^d . (The Haar measure is the unique nonzero measure which is invariant under the group operation.)

\mathbb{P} (which is known as *twirling*) takes arbitrary density matrices $A \in \mathcal{D}(\mathbb{C}^d \otimes \mathbb{C}^d)$ and returns Werner states. Indeed, due to the invariance of dU and the fact that product of

unitaries is also unitary, it follows

$$\begin{aligned}
\tilde{U} \otimes \tilde{U} \mathbb{P}(A) \tilde{U}^\dagger \otimes \tilde{U}^\dagger &= \int \tilde{U} \otimes \tilde{U} (U \otimes U) A (U^\dagger \otimes U^\dagger) \tilde{U}^\dagger \otimes \tilde{U}^\dagger dU \\
&= \int (\tilde{U} U \otimes \tilde{U} U) A (U^\dagger \tilde{U}^\dagger \otimes U^\dagger \tilde{U}^\dagger) dU \\
&= \int (U \otimes U) A (U^\dagger \otimes U^\dagger) dU \\
&= \mathbb{P}(A),
\end{aligned}$$

which also proves that \mathbb{P} is a projection (onto the $U \otimes U$ -invariant operators subspace). Then $\mathbb{P}(A)$ depends only on the parameter $\text{tr}(V\mathbb{P}(A))$, and we can utilize the $U \otimes U$ -invariance of V to show that

$$\begin{aligned}
\text{tr}(V\mathbb{P}(A)) &= \int \text{tr}(V(U \otimes U) A (U^\dagger \otimes U^\dagger)) dU = \\
&= \int \text{tr}((U^\dagger \otimes U^\dagger) V (U \otimes U) A) dU = \text{tr}(VA).
\end{aligned}$$

Also, if A is separable, so is each $(U \otimes U) A (U^\dagger \otimes U^\dagger)$ and hence $\mathbb{P}(A)$ is also separable. This means that \mathbb{P} preserves the separability and the trace of the operator times V .

Now, let $\rho = \rho_0 \otimes \rho_1, \sigma = \sigma_0 \otimes \sigma_1 \in \mathcal{D}(\mathbb{C}^d \otimes \mathbb{C}^d)$ be separable states. Taking ρ_0 and ρ_1 to be orthogonal (e.g.: $\rho_0 = |0\rangle\langle 0|, \rho_1 = |1\rangle\langle 1|$) and making use of Prop. 3.4.1, we have

$$\text{tr}(V\mathbb{P}(\rho)) = \text{tr}(V\rho) = \text{tr}(\rho_0\rho_1) = 0.$$

Thus $\mathbb{P}(\rho)$ is precisely the Werner state determined by $\phi = 0$, which is W_0 . Since \mathbb{P} preserves separability, we conclude that W_0 is separable.

Similarly, for $\sigma_0 = \sigma_1 = |\psi\rangle\langle\psi|$ we have

$$\text{tr}(V\mathbb{P}(\sigma)) = \text{tr}(V\sigma) = \text{tr}(|\psi\rangle\langle\psi|) = 1.$$

Thus $\mathbb{P}(\sigma) = W_1$, whose separability is implied by the separability of σ , completing the proof. \square

The first part of the proof establishes that the flip operator is an entanglement witness² for quantum states, i.e., we can calculate $\text{tr}(V\rho)$ for any state ρ and conclude that it is entangled if the result is negative (although, in general, no conclusion can be taken if the trace results to be positive). The second part shows that the flip is the optimal witness for Werner states, since it reveals the entanglement of any of such states.

²It seems fair to say that the concept of entanglement witness only appeared in Ref. [23], dating from 1996 - hence, posterior to Werner's paper.

4.2 Bob's response function and the hidden variables

Now that we have setted the class of states, we focus on the objects that compose the local model: the hidden variables space, the probability distribution upon it and the response functions. The consequences of the $U \otimes U$ -invariance of Werner states is a key observation to motivate the choices we are going to make.

Since W is $U \otimes U$ -invariant, we have that the left-hand side of Eq. (3.7) is

$$\begin{aligned} \text{tr}(W P_a \otimes Q_b) &= \text{tr}(U \otimes U W U^\dagger \otimes U^\dagger P_a \otimes Q_b) \\ &= \text{tr}(W U^\dagger \otimes U^\dagger P_a \otimes Q_b U \otimes U) = \text{tr}(W (U^\dagger P_a U) \otimes (U^\dagger Q_b U)). \end{aligned}$$

Since $(U^\dagger P_i U)^2 = U^\dagger P_i U$ and $\sum_i U^\dagger P_i U = U^\dagger \sum_i P_i U = I$, we see that $\{U^\dagger P_i U\}$ also defines a projective measurement. Thus the symmetry of the Werner states implies that the probabilities obtained for any pair of measurements in $\{(\{U P_i U^\dagger\}, \{U Q_j U^\dagger\}) \mid U \in \mathcal{L}(\mathcal{H}) \text{ unitary}\}$ must be the same. Therefore, the local model to be constructed must satisfy

$$\int_{\Lambda} p_A(a, \lambda, \{P_i\}) p_B(b, \lambda, \{Q_j\}) \pi(\lambda) d\lambda = \int_{\Lambda} p_A(a, \lambda, \{U P_i U^\dagger\}) p_B(b, \lambda, \{U Q_j U^\dagger\}) \pi(\lambda) d\lambda. \quad (4.4)$$

This motivates the setting of the following arrangements, in the construction of the local model.

The space Λ of hidden variables will be chosen to be the unit sphere $\{\lambda \in \mathbb{C}^d; \|\lambda\| = 1\}$, that is, the hidden variables can be seen as pure quantum states of the local systems. We will also impose a symmetry condition to the response functions p_A and p_B used by Alice and Bob. These will be taken to depend on the family $\{P_i\}$ of orthogonal projections, but not on the their labeling or their eigenvalues. Moreover, it suffices to consider only the case where the projectors are one-dimensional, since for projections of higher dimension the response functions can be chosen as a sum of response functions of one-dimensional projections. The symmetry condition to be imposed on the response functions is given by the relation

$$p(a, \lambda, \{U^\dagger P_i U\}) = p(a, U\lambda, \{P_i\}). \quad (4.5)$$

With this two requirements, Eq. (4.4) is equivalent to

$$\int_{\Lambda} p_A(a, \lambda, \{P_i\}) p_B(b, \lambda, \{Q_j\}) \pi(\lambda) d\lambda = \int_{\Lambda} p_A(a, U\lambda, \{P_i\}) p_B(b, U\lambda, \{Q_j\}) \pi(\lambda) d\lambda,$$

which is satisfied if we set the measure $\pi(\lambda)$ to be the unique measure invariant under unitaries on Λ .

Hence, to complete our description of the local model, we still are left to decide the

response functions p_A and p_B , respecting Eq. (4.5). The simplest choice for such functions is the one we will adopt for Bob's system,

$$p_B(b, \lambda, \{Q_j\}) = \text{tr}(Q_b |\lambda\rangle \langle \lambda|) = \langle \lambda | Q_b | \lambda \rangle, \quad (4.6)$$

for the observable $Q = \sum_b \beta_b Q_b$. Notice that, by doing this, we are saying that Bob does things “in the Quantum fashion”, since his probabilities will be given by treating the hidden variable as a pure state and using Born's rule to calculate probabilities. Also, the function works as we wish in (4.5),³

$$p_B(b, \lambda, \{U^\dagger Q_j U\}) = \langle \lambda | U^\dagger Q_b U | \lambda \rangle = \langle U \lambda | Q_b | U \lambda \rangle = p_B(b, U \lambda, \{Q_j\}).$$

We now proceed to show that using Eqs. (4.5) e (4.6), we will be able to reduce the calculation of any of the integrals in the form of Eq. (3.7) to the computation of a single integral.

For any positive integrable function $f : \Lambda \rightarrow \mathbb{R}$, we can consider the positive operator⁴

$$\tilde{f} = \int_{\Lambda} |\lambda\rangle \langle \lambda| f(\lambda) \pi(\lambda) d\lambda \in \mathcal{L}(\mathbb{C}). \quad (4.7)$$

Then, using Bob's response function with $\{Q_j\}$ and b fixed, we have

$$\int_{\Lambda} p_B(b, \lambda, \{Q_j\}) f(\lambda) \pi(\lambda) d\lambda = \int_{\Lambda} \text{tr}(Q_b |\lambda\rangle \langle \lambda|) f(\lambda) \pi(\lambda) d\lambda = \text{tr}(Q_b \tilde{f}).$$

In particular, for each fixed $\{P_i\}$ and a , we can relate to p_A the positive operator \tilde{p}_A given in Eq. (4.7), such that

$$\int_{\Lambda} p_B(b, \lambda, \{Q_j\}) p_A(a, \lambda, \{P_i\}) \pi(\lambda) d\lambda = \text{tr}(\tilde{p}_A(a, \{P_i\}) Q_b). \quad (4.8)$$

Notice that the left-hand side of the above equation equals the right-hand side of Eq.

³For a comment on the abuse of Dirac's notation done here, check Section 1.1.

⁴There is a misprint in the definition of $\tilde{\rho}$ in the original paper: the $f(\lambda)$ is missing, in the integrand.

(3.7). Using the U -invariance of π , we find

$$\begin{aligned}
\text{tr}(U\tilde{p}_A(a, \{P_i\})U^\dagger Q_b) &= \text{tr}(\tilde{p}_A(a, \{P_i\})U^\dagger Q_b U) \\
&= \int_{\Lambda} p_A(a, \lambda, \{P_i\}) \langle \lambda | U^\dagger Q_b U | \lambda \rangle \pi(\lambda)(d\lambda) \\
&= \int_{\Lambda} p_A(a, U\lambda, \{P_i\}) \langle \lambda | Q_b | \lambda \rangle \pi(\lambda)(d\lambda) \\
&= \int_{\Lambda} p_A(a, \lambda, \{UP_iU^\dagger\}) \langle \lambda | Q_b | \lambda \rangle \pi(\lambda)(d\lambda) \\
&= \text{tr}(\tilde{p}_A(a, \{UP_iU^\dagger\})Q_b).
\end{aligned}$$

Since this holds for all one-dimensional projections Q_b , we have

$$U\tilde{p}_A(a, \{P_i\})U^\dagger = \tilde{p}_A(a, \{UP_iU^\dagger\}).$$

In particular, we see that, if U commutes with every P_i , then U commutes with \tilde{p}_A :

$$U\tilde{p}_A(a, \{P_i\})U^\dagger = \tilde{p}_A(a, \{UP_iU^\dagger\}) = \tilde{p}_A(a, \{P_iUU^\dagger\}) = \tilde{p}_A(a, \{P_i\}).$$

That is to say (check Corollary 1) that \tilde{p}_A (and the same is true for U) has a representation

$$\tilde{p}_A(a, \{P_i\}) = \sum_i \omega(a, i) P_i.$$

Since p_A is not to depend on the labeling of the projections P_i , we conclude that $\omega(a, i)$ depends only on whether $a = i$ or not. Hence

$$\tilde{p}_A(a, \{P_i\}) = \omega P_a + \omega' \sum_{i \neq a} P_i = \omega_1 P_a + \omega_2 I,$$

for some $\omega_1, \omega_2 \in \mathbb{R}$, which are independent of P_i , $i \neq a$. Since

$$\begin{aligned}
\sum_a \tilde{p}_A(a, \{P_i\}) &= \sum_a \int_{\Lambda} p_A(a, \lambda, \{P_i\}) |\lambda\rangle \langle \lambda| \pi(\lambda) d\lambda \\
&= \int_{\Lambda} \sum_a (p_A(a, \lambda, \{P_i\})) |\lambda\rangle \langle \lambda| \pi(\lambda) d\lambda \\
&= \int_{\Lambda} |\lambda\rangle \langle \lambda| \pi(\lambda) d\lambda \\
&= I
\end{aligned}$$

we must have

$$\sum_a (\omega_1 P_a + \omega_2 I) = \omega_1 I + d\omega_2 I = I,$$

thus

$$\omega_1 + d\omega_2 = 1. \quad (4.9)$$

Hence for computing $\tilde{p}_A(a, \{P_i\})$ for a given a , it suffices to calculate the value of the expression $\text{tr}(Q_b \tilde{p}_A(a, \{P_i\}))$ for only one arbitrary choice of Q_b . (This is the great advantage of dealing with \tilde{p}_A : it is uniform over all Q_b .) Choosing $Q_b = P_a$ [1], we have

$$\text{tr}(\tilde{p}_A(a, \{P_i\})P_a) = \text{tr}([\omega_1 P_a + \omega_2 I]P_a) = \text{tr}(\omega_1 P_a + \omega_2 P_a) = \omega_1 + \omega_2. \quad (4.10)$$

Assuming in advance that our local model works, i.e., that

$$\text{tr}(W P_a \otimes Q_b) = \int_{\Lambda} p_A(a, \lambda, \{P_i\}) p_B(b, \lambda, \{Q_j\}) \pi(\lambda) d\lambda,$$

and using Eq. (4.8), we have

$$\text{tr}(W P_a \otimes Q_b) = \text{tr}(Q_b \tilde{p}_A(a, \{P_i\})),$$

hence $\text{tr}(Q_b \tilde{p}_A(a, \{P_i\}))$ can be easily calculated (once W, P_a and Q_b are determined and well-known) and Eqs. (4.9), (4.10) defines explicitly ω_1, ω_2 and thus \tilde{p}_A .

However, it is not really important the values of ω_1, ω_2 . The main point here is that, for fixed $\{P_i\}$ and a , solving the problem for $Q_b = P_a$ implies solving it for any Q_b . In this situation, the only probability we need to achieve with our local model is

$$\begin{aligned} \text{tr}(W P_a \otimes P_a) &= \frac{(d - \phi)(\text{tr } P_a)(\text{tr } P_a) + (d\phi - 1) \text{tr}(P_a^2)}{d^3 - d} \\ &= \frac{(d - \phi) + (d\phi - 1)}{d^3 - d} \\ &= \frac{d(1 + \phi) - (1 + \phi)}{d^3 - d} \\ &= \frac{(d - 1)(1 + \phi)}{d(d^2 - 1)} \\ &= \frac{1 + \phi}{d(d + 1)}, \end{aligned}$$

that is, it suffices to show that

$$\frac{1 + \phi}{d(d + 1)} = \int_{\Lambda} p_A(a, \lambda, \{P_i\}) \langle \lambda | P_a | \lambda \rangle \pi(\lambda) d\lambda.$$

But this last equality is trivially achieved by setting

$$\phi = -1 + d(d + 1) \int_{\Lambda} p_A(a, \lambda, \{P_i\}) \langle \lambda | P_a | \lambda \rangle \pi(\lambda) d\lambda. \quad (4.11)$$

In other words, we have constructed a local model for the state given by the parameter ϕ determined by Eq. (4.11). However, we do not have much information about this state. Particularly, we are still under the risk of the integral in Eq. (4.11) provides a parameter ϕ which determines a separable state, and thus we are only presenting a sophisticated proof of a well known fact to us: that such separable state is local.

Therefore, it remains to show that the right-hand side of Eq. (4.11) can be negative, what would gives us a negative ϕ , corresponding (according with Theorem 6) to an entangled state. In order to do that, the card still left up our sleeve is the setting of Alice's response function $p_A(a, \lambda, \{P_i\})$. Our task now is to determine p_A satisfying the symmetry (4.5) for which the integral in (4.11) becomes as small as possible (since it is always nonnegative), under the constraints $p_A(a, \lambda, \{P_i\}) \geq 0$ and $\sum_a p_A(a, \lambda, \{P_i\}) = 1$ for all λ and all $\{P_i\}$.

We observe that Eq. (4.11) determines ϕ (and thus W) according with the choice of Alice's measurement $\{P_i\}$ and outcome a . So, in principle, it seems that we have constructed a state that depends on the measurement to be carried out on Alice's side. This would be a huge problem, since our goal is, for a given a state, be able to simulate the correlations provided by *any* local measurements. We will see, however, that the choice of Alice's response function and the measure $\pi(\lambda)$ guarantees that the integral in (4.11) equals the same value for any $a, \{P_i\}$ under consideration.

4.3 Alice's response function and a local model for an entangled state

At this point, we might feel tempted to look for a manner to compute the probabilities of Alice's outcomes similar to Bob's. The response function $p_B(b, \lambda, \{Q_j\})$ employed by Bob has the characteristic property of depending only on Q_b , once b and λ are fixed. The remaining measurement operators do not influence p_B . This may appear very natural, since Born's rule got us used to something like this, in the quantum context. Nevertheless, it is vital for Werner's construction that the response function p_A does not have this property.

Indeed, suppose that, by fixing λ , we have $p_A = p_A(P_a)$ (as it is p_B). Then p_A is a non-negative, summing to one and additive map on families of mutually orthogonal projections, while response function. Associated to p_A , there is a probability measure μ such that, if P_i is the projector onto the subspace $E_i \subset \mathbb{C}^d$, then

$$p_A(P_i) = \mu(E_i),$$

in the sense of Def. 2.6. According to Gleason's theorem (and assuming $d > 2$), there is a density operator ρ_λ such that

$$p_A(P_a) = \text{tr}(P_a \rho_\lambda).$$

Substituting such p_A as well as the previously defined p_B in the left hand side of Eq. (3.7), we have

$$\int_{\Lambda} \text{tr}(P_a \rho_\lambda) \text{tr}(Q_b |\lambda\rangle \langle \lambda|) \pi(\lambda) d\lambda = \int_{\Lambda} \text{tr}(P_a \otimes Q_b \rho_\lambda \otimes |\lambda\rangle \langle \lambda|) \pi(\lambda) d\lambda.$$

Hence, those response functions give rise to a local model that simulates the mixed state

$$\int_{\Lambda} \rho_\lambda \otimes |\lambda\rangle \langle \lambda| \pi(\lambda) d\lambda.$$

However, the above state is clearly separable and, thus, local *a priori*! (Remember, the objective of our quest is an *entangled* local state.) Therefore, we conclude that the response function p_A should depend on other projectors of $\{P_i\}$ and not just on P_a .

With this in our minds, there is only one more observation to be made before the setting of p_A . Since for every fixed λ and P the constraints

$$p_A(a, \lambda, \{P_i\}) \geq 0, \quad \sum_a p_A(a, \lambda, \{P_i\}) = 1$$

single out a convex set in \mathbb{R}^d , we expect that the smallest values of ϕ is attained for response functions taking only the values 0 and 1. This suggests the following choice:

$$p_A(a, \lambda, \{P_i\}) = \begin{cases} 1, & \text{if } \forall k \neq a, \langle \lambda | P_a | \lambda \rangle < \langle \lambda | P_k | \lambda \rangle \\ 0, & \text{if } \exists k \neq a; \langle \lambda | P_a | \lambda \rangle > \langle \lambda | P_k | \lambda \rangle. \end{cases} \quad (4.12)$$

Note that we have left $p_A(a, \lambda, \{P_i\})$ unspecified at all points where $\langle \lambda | P_a | \lambda \rangle$ is the minimum of $\langle \lambda | P_k | \lambda \rangle$ but not the unique one. However, since this set is of measure zero, it will not contribute to the integral (4.11) anyway, and we may choose on this subset any measurable function satisfying the constraint.

We have written Eq. (4.12) in such a form that the property postulated in Eq. (4.5) is manifest. Moreover, p_A is independent of the labeling of the P_a in the sense that it only depends of the set of numbers $\langle \lambda | P_k | \lambda \rangle$, but not on their ordering.

Substituting (4.12) in the integrand in (4.11), we have

$$\int_{\Lambda} p_A(a, \lambda, \{P_i\}) \langle \lambda | P_a | \lambda \rangle \pi(\lambda) d\lambda = \int_{\Lambda_a} \langle \lambda | P_a | \lambda \rangle \pi(\lambda) d\lambda, \quad (4.13)$$

where $\Lambda_a = \{\lambda \in \Lambda; \forall k \neq a, \langle \lambda | P_a | \lambda \rangle < \langle \lambda | P_k | \lambda \rangle\}$. Setting $\langle \lambda | P_i | \lambda \rangle = u_i$, through the relation

$$|\lambda\rangle \mapsto (u_1, \dots, u_d)$$

we can identify the space of pure states of $\mathbb{C}^d \otimes \mathbb{C}^d$ with the simplex of d vertices $\mathcal{S} = \{(u_1, \dots, u_d) \in \mathbb{R}^d; u_1 + \dots + u_d = 1\}$, which is embedded in \mathbb{R}^{d-1} , and Λ_a with the subset $\mathcal{S}' \subset \mathcal{S}$ delimited by the hyperplanes $u_a = u_1, \dots, u_a = u_d$. The simplex is best imagined to “stand” upon the plane $u_a = 0$, so that the u_a represents the “vertical” axis. \mathcal{S}' thus is the convex set formed by the face of \mathcal{S} resting on the plane $u_a = 0$ and vertex on the barycenter of \mathcal{S} , which is at height $u_a = 1/d$.

To illustrate the above reasoning, let's calculate explicitly the integral in the case where $d = 3$. The simplex turns out to be the restricted plane $\mathcal{S} = \{(x, y, z) \in \mathbb{R}^3; x + y + z = 1, x, y, z \geq 0\}$ and the planar set $\mathcal{S}' \subset \mathcal{S}$ is delimited by the lines $x = y$, $2z = 1 - x$ and $2z = 1 - y$.

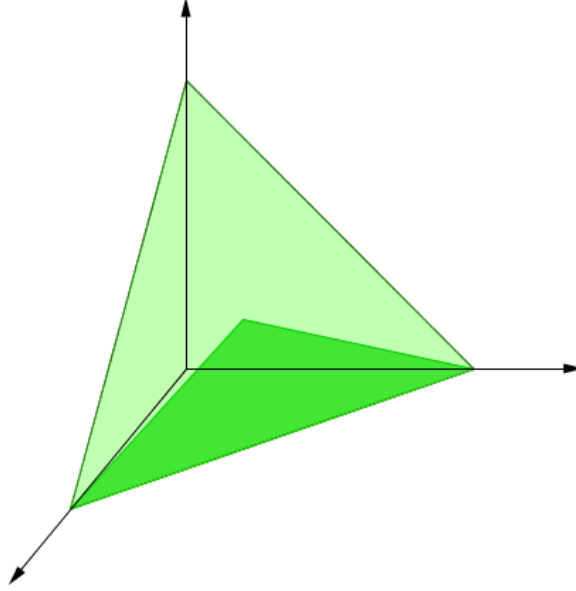


Figure 4.1: The simplex $\mathcal{S} \subset \mathbb{R}^3$ in light color and the subset \mathcal{S}' in dark.

The barycenter of \mathcal{S} (and apex of \mathcal{S}') is the point $(1/3, 1/3, 1/3)$. We can parameterize the plane containing the simplex using

$$f(r, s) = (r, s, 1 - r - s), \quad r, s \in \mathbb{R},$$

for which we have

$$f_r(r, s) = (1, 0, -1), \quad f_s(r, s) = (0, 1, -1), \quad \|f_r \wedge f_s\| = \sqrt{3}.$$

We want to integrate the function $g(r, s) = 1 - r - s$ over the domain comprehended

between the points $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 1/3, 1/3)$. Hence, the integral is given by

$$\begin{aligned} & \int_0^{\frac{1}{3}} \int_{-2x+1}^{-x+1} (1-r-s)\sqrt{3}drds + \int_{\frac{1}{3}}^1 \int_{-x/2+1/2}^{-x+1} (1-r-s)\sqrt{3}drds \\ &= \frac{\sqrt{3}}{2 \cdot 3^4} + \frac{\sqrt{3}}{3^4} \\ &= \frac{\sqrt{3}}{2} \frac{1}{3^3}. \end{aligned}$$

The normalization (indicated by $\pi(\lambda)$ in Eq. (4.13)) is done by dividing the resulting expression by $\sqrt{3}/2$, the area of \mathcal{S} , an equilateral triangle of side length $\sqrt{2}$. We obtain the result $1/3^3$.

In the general case, the integral of u_a over \mathcal{S}' is equal to the height of the barycenter of \mathcal{S}' times the volume of \mathcal{S}' . Since the barycenter has height $u_a = 1/d^2$ and the volume of \mathcal{S}' is $1/d$ (the simplex is formed by d pieces congruent to \mathcal{S}' and has volume equal to 1 by definition), we find that the integral results in $1/d^3$.

Substituting this value in Eq. (4.11), we have

$$\phi = -1 + \frac{1+d}{d^2}. \quad (4.14)$$

Henceforth, we have constructed a local model for projective measurements for an entangled Werner state, since this value of ϕ is negative for all $d \geq 2$. We will denote this state by W_{local} .

By now, after the considerations over p_A and the correspondence with the simplex, it is clear that the value of ϕ in Eq. (4.11) is independent of the choice of P_a and the whole construction results in the state given by 4.14 regardless of our starting point.

4.4 Barrett's local model

Werner's paper [1] was a breakthrough in the foundations of Quantum Theory⁵, since showed that entanglement and non-locality are distinct features of quantum states. His model is a proof that some aspects of Quantum Theory can be reproduced using only classical resources and that nonlocality is one of the distinctly nonclassical features of Quantum Theory.

However, the reproduction capability presented by Werner's local model is limited, since it only works for projective measurements. Indeed, quantum mechanics allows us to perform a more general kind of measurement, the POVMs. The natural question raised now is: is there a local model which can reproduce the correlations generated by any

⁵However, it took a while for being acclaimed as so. It is nice to see the history of citations.

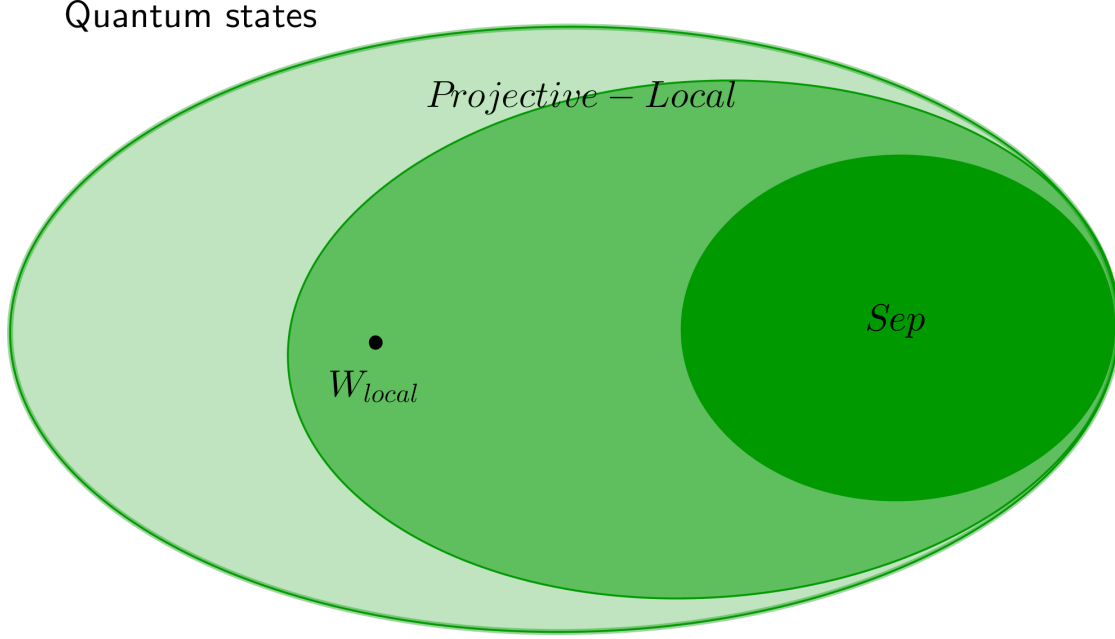


Figure 4.2: Werner’s local model shows that the set of separable states is strictly contained in the set of projective-local states: at least W_{local} is projective-local but not separable.

POVM applied to an entangled state?

Jonathan Barrett showed that the answer to this question is affirmative [25]. Choosing appropriate response functions for Alice and Bob, he was able to create a model that reproduces the correlations given by any POVM over an slightly modified Werner state which is entangled. More explicitly, the state simulated by Barrett is

$$W_{Barrett} = \alpha \frac{2 \sum_{i < j; i, j=1}^d |ij\rangle \langle ji|}{d(d-1)} + (1 - \alpha) \frac{I_{d \times d}}{d^2}, \quad (4.15)$$

where

$$\alpha = \frac{(d-1)^{d-1}(3d-1)}{(d+1)d^d}$$

and d is the local dimension. (In Chapter 5 we show that the original Werner state also has the form in Eq. (4.15), with $\alpha = (d-1)/d$.) In state $W_{Barrett}$, the first density matrix that appears in the superposition is sometimes called the *projector onto the antisymmetric subspace* of $\mathcal{D}(\mathbb{C}^d \otimes \mathbb{C}^d)$. $W_{Barrett}$ can be shown to be entangled if and only if $\alpha > 1/(1+d)$, which happens for all $d \geq 2$.

Quantum states

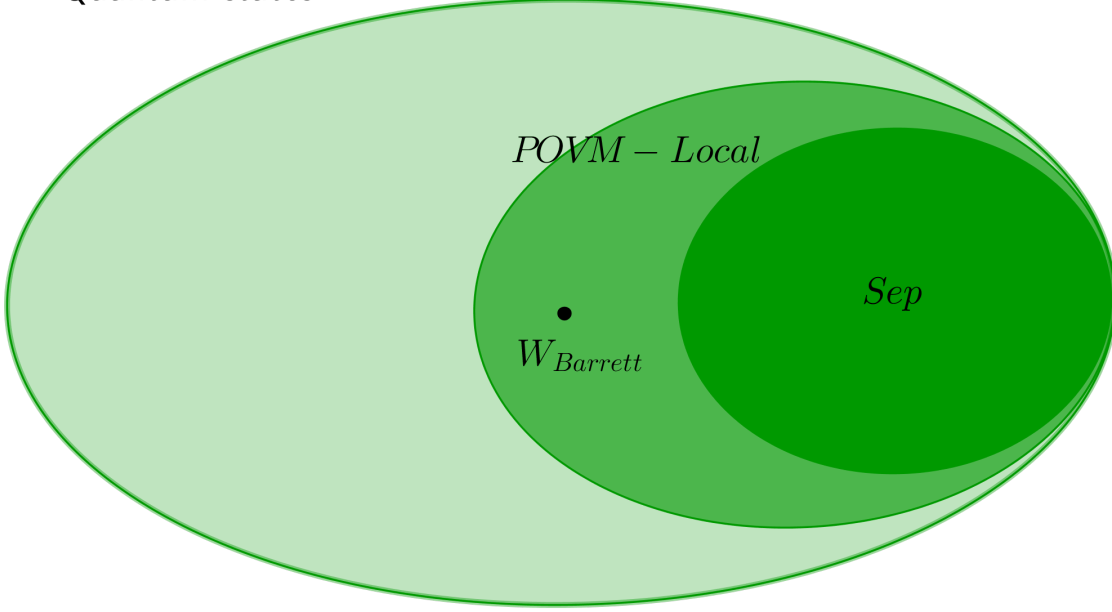


Figure 4.3: Barrett’s local model shows that the set of separable states is strictly contained in the set of POVM-local states: at least $W_{Barrett}$ is POVM-local but not separable. It is still an open question if the sets of projective-local and POVM-local states are the same or not; in other words, it is not known if general measurements indeed offer any advantage to detect the nonlocality of projective-local states.

One of the key observations that allowed Barrett to construct the model was that we can restrict our attentions to the case where all the POVM elements are proportional to one-dimensional projectors, that is, POVMs of the form $M = \{M_i\}$ where $M_i = a_i P_i$, with $0 \leq a_i \leq 1$ and P_i one-dimensional projector. Though we are not going to detail Barrett’s local model, this simple result will be further used in Chapter 7, where another local model for POVM will be detailed, so we prove it now.

Proposition 4.4.1. *If a local model simulates the correlations of a POVM of the form $M = \{M_i\}$, where $M_i = a_i P_i$, $0 \leq a_i \leq 1$ and P_i is a rank-1 projector, then it simulates the correlations of any POVM.*

Proof. Let $N = \{N_j\}$ be an arbitrary POVM. According to the Spectral Decomposition Theorem (Theorem 2), each positive operator N_j can be written as $\sum_k c_{jk} P_{jk}$, where $0 \leq c_{jk} \leq 1$ and the P_{jk} are one-dimensional operators satisfying $P_{jk} P_{jk'} = \delta_{kk'} P_{jk}$. If each N_j is written in this form, then we can regard as if a more “detailed” POVM is taking

place than the one who is actually been performed, with elements $\{c_{jk}P_{jk}\}$, and our model will make the appropriate predictions. If the outcome referent to P_{jk} is predicted by the model, then we say that the outcome N_j was actually obtained. \square

In Barrett's model, the set of hidden variables is the same as the one used by Werner, tridimensional unit vectors $\lambda \in \mathbb{C}^d$, considered with uniform distribution. The response functions defined for the POVMs $M = \{M_i\} = \{x_i P_i\}$ performed by Alice is

$$\begin{aligned} p_A(i, \lambda, \{M_i\}) &= \langle \lambda | M_i | \lambda \rangle \chi_{\{x \in \mathbb{R}; x \geq 0\}} (\langle \lambda | P_i | \lambda \rangle - 1/d) \\ &\quad + \left(1 - \sum_k \langle \lambda | M_k | \lambda \rangle \chi_{\{x \in \mathbb{R}; x \geq 0\}} (\langle \lambda | P_k | \lambda \rangle - 1/d) \right) \frac{x_i}{d}, \end{aligned}$$

where the characteristic function χ_S is defined to map s to 1 if $s \in S$ and to 0 otherwise. Bob's response function for $N = \{N_j\} = \{y_j Q_j\}$ is

$$p_B(j, \lambda, \{N_j\}) = \frac{1}{d-1} y_j (1 - \langle \lambda | Q_j | \lambda \rangle).$$

More about this subject will be said in Chapter 7.

Chapter 5

Hidden nonlocality

Until Sandu Popescu's work [3], all the treatment of Bell inequalities had a common aspect: they all consider the case in which the system is subjected to a *single* local measurement in each part. Popescu showed that, despite the Werner state referent to the parameter ϕ in Eq. (4.14) being local for any single measurement, i.e., do not violate any Bell inequality, for a *sequence of measurements*, it does. This property became known as *hidden nonlocality*.

The local state provided by Werner is

$$W_{local} = \frac{d+1}{d^3} I_{d \times d} - \frac{1}{d^2} V. \quad (5.1)$$

Defining $|S_{ij}\rangle$ to be “the singlet state in positions i and j ”, that is,

$$|S_{ij}\rangle = \frac{1}{\sqrt{2}}(|ij\rangle - |ji\rangle),$$

we have

$$-2 |S_{ij}\rangle \langle S_{ij}| = |ij\rangle \langle ji| + |ji\rangle \langle ij| - |ij\rangle \langle ij| - |ji\rangle \langle ji|,$$

thus the flip operator V can be written as

$$V = I - 2 \sum_{i < j; i, j=1}^d |S_{ij}\rangle \langle S_{ij}|.$$

Substituting the above relation in Eq. (5.1), we can write the Werner state as

$$W = \frac{1}{d^2} \left(\frac{1}{d} I + 2 \sum_{i < j; i, j=1}^d |S_{ij}\rangle \langle S_{ij}| \right). \quad (5.2)$$

We will now show that after Alice and Bob perform each one a large rank projective

measurement, it is possible that the resulting state violates the CHSH inequality for specified measurements A, A', B and B' .

First, each part performs the measurement referent to the projective operator

$$P = |1\rangle\langle 1| + |2\rangle\langle 2|,$$

that is, the measurement elements involved are $\{P, I - P\}$. The resulting state belongs to the space generated by $\{|1\rangle, |2\rangle\}$ (in which case we will say the outcome was 1) or to the space generated by $\{|3\rangle, \dots, |d\rangle\}$ (in which case we will say the outcome was 0).

Because each part is performing this measurement, we have four possible outcomes: $(0, 0), (0, 1), (1, 0)$ and $(1, 1)$. The resulting (unnormalized) state corresponding to the outcome $\{1, 1\}$ is

$$\begin{aligned} P \otimes P W_{local} P \otimes P &= \frac{1}{d^3} P \otimes P I P \otimes P + \frac{2}{d^2} \sum_{i < j} P \otimes P |S_{ij}\rangle \langle S_{ij}| P \otimes P \\ &= \frac{1}{d^3} P \otimes P + \frac{2}{d^2} |S_{12}\rangle \langle S_{12}|, \end{aligned}$$

where we have used the form (5.2). Let's denote the subspace generated by $\{|1\rangle, |2\rangle\}$ by \mathcal{S}_{12} . Notice that $P \otimes P := I^{4 \times 4}$ is the 4×4 identity matrix acting in the $\mathcal{S}_{12} \otimes \mathcal{S}_{12}$ subspace and zero at the rest. Recalling that $|S_{12}\rangle = |\Psi_{-}\rangle$, after normalization we obtain

$$\begin{aligned} W' &= \frac{1}{\text{tr}(\frac{1}{d^3} I^{4 \times 4} + \frac{2}{d^2} |\Psi_{-}\rangle \langle \Psi_{-}|)} \left(\frac{1}{d^3} I^{4 \times 4} + \frac{2}{d^2} |\Psi_{-}\rangle \langle \Psi_{-}| \right) \\ &= \frac{1}{(\frac{4}{d^3} + \frac{2}{d^2}) d^3} I^{4 \times 4} + \frac{2}{(\frac{4}{d^3} + \frac{2}{d^2}) d^2} |\Psi_{-}\rangle \langle \Psi_{-}| \\ &= \frac{d}{d+2} \left(\frac{1}{2d} I^{4 \times 4} + |\Psi_{-}\rangle \langle \Psi_{-}| \right). \end{aligned}$$

Notice that as d grows, the state approaches the singlet.

For the second measurement, Alice chooses¹ randomly between operators A and A' and Bob between B and B' . Each of these operators have three different eigenvalues, 1, -1 and 0. The eigenvalues 1 and -1 are non-degenerate and the corresponding eigenvectors belong to the subspace \mathcal{S}_{12} . The eigenvalue 0 is highly degenerate and corresponds to the rest of the space, the subspace generated by $\{|3\rangle, \dots, |d\rangle\}$. The nongenerate part of these operators is chosen such that they yield maximal violation of the CHSH inequality for the singlet state $|\Psi_{-}\rangle$, that is,

$$\langle \Psi_{-} | AB + AB' + A'B - A'B' | \Psi_{-} \rangle = 2\sqrt{2}.$$

¹At this point, perhaps should be made a comment on a delicate topic, "causality". We won't.

In another words, the operators A, A', B, B' , when restricted to $\mathcal{S}_{12} \otimes \mathcal{S}_{12}$, are equal to the operators Q, R, S, T , respectively, defined in Section 3.1.

With these operators and the state W' , we have

$$\begin{aligned} & \text{tr}(W'[AB + AB' + A'B - A'B']) \\ &= \frac{1}{2d+4} \text{tr}(QS + QT + RS - RT) + \frac{d}{d+2} \langle \Psi_- | AB + AB' + A'B - A'B' | \Psi_- \rangle. \end{aligned}$$

It is not hard to see that $\text{tr}(QS) = \text{tr}(RS) = \text{tr}(RT) = \text{tr}(QT) = 0$, hence the first term of the right side of the above equation vanishes and we conclude

$$\text{tr}(W'[AB + AB' + A'B - A'B']) = \frac{2\sqrt{2}d}{d+2} > 2$$

for $d > 5$.

We conclude that, although Werner's model can simulate all the correlations which arise when only a single measurement is performed on each of the two particles, the model cannot account for the correlations which arise when two consecutive measurements are performed in each particle.

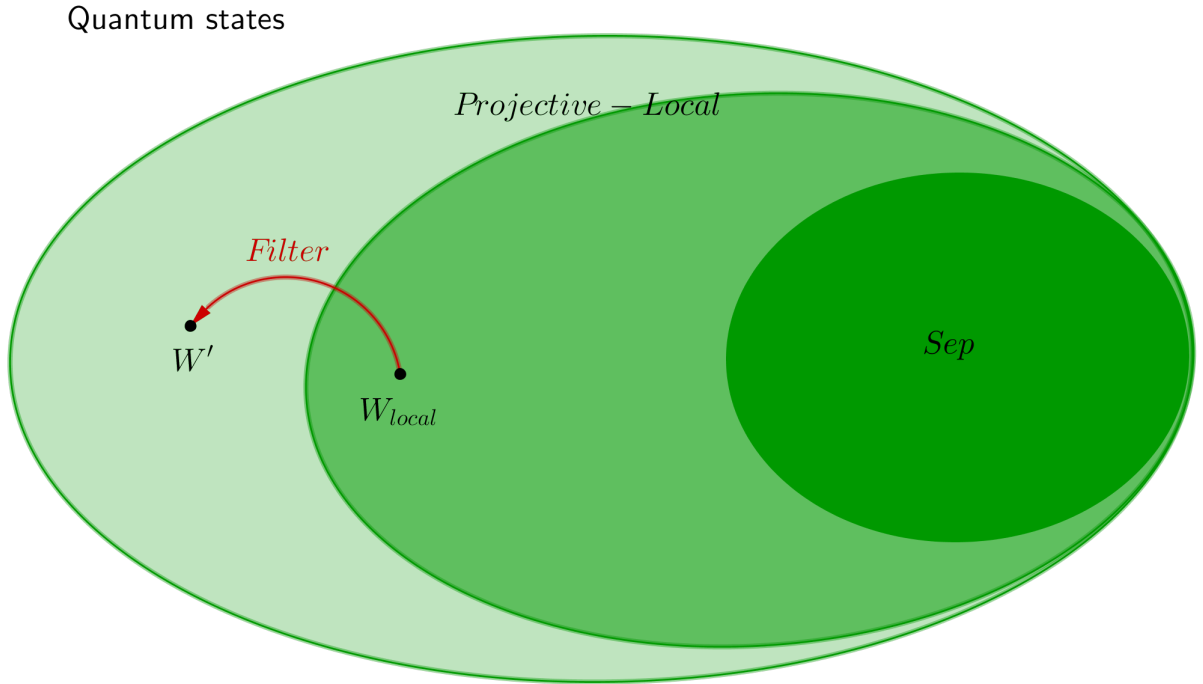


Figure 5.1: The successful application of a local filtering (represented by the arrow in red) on W_{local} results in a state W' which violates CHSH, and thus does not belong to the set of local states.

The main idea here is the introduction of previous local measurements, which are able

to take the Werner state W_{local} to a CHSH-violating state with non-zero probability. The nonlocality revelation depends on the success of these measurements. Therefore, they are known as *local filters*: they filter the protocol of revealing nonlocality, in the sense that the process is discarded if they fail, not providing the desired state needed for a Bell inequality violation. In Chapter 7 we will meet such objects again.

Chapter 6

Gisin-Degorre's local model

In 1999, Nicolas Gisin and Bernard Gisin presented a local model for the singlet state based on the detection loophole, that is, the non-maximal efficiency inherent to detectors, in experiments [8]. The idea was to introduce a “probability of failure” in the detector of one system (say Alice’s), generated by the local hidden variables.

Years later, in 2005, Julien Degorre and co-workers studied a way to quantify nonlocality by measuring the amount of additional resources that should be considered besides the hidden variables in order to simulate the measurements’ correlations, namely classical communications, post-selection and nonlocal boxes [9]. They considered the Gisin-Gisin model in [8] and reinterpreted the probability of failure as a probability of rejection by Alice. Then, with only one bit of communication, they were able to reproduce locally the statistics of the outcomes obtained in a EPR experiment (see Section 2.5). However, without any further resources, their resulting protocol happens to simulate the Werner state of $\mathbb{C}^2 \otimes \mathbb{C}^2$

$$W = \alpha |\psi_-\rangle \langle \psi_-| + (1 - \alpha) \frac{I_{4 \times 4}}{4},$$

with parameter $\alpha = 1/2$.

In order to agree in notation with the original paper [9], from this point on we are going to use the scalar product symbol for the functional of $\mathbb{R}^3 \times \mathbb{R}^3$, which for unit vectors denotes the co-sine of the angle between them, i.e.,

$$v \cdot u := \cos \theta_{vu},$$

for unit vectors $|v\rangle, |u\rangle \in \mathbb{R}^3$.

6.1 Shared randomness

The context of [9] is Information Theory, and this influences their approach to the local models problem. The first important observation is that the local hidden variables can be seen as shared randomness, that is, as random variables provided by some source to which both Alice and Bob have access. As discussed in Chapter 3, the role of the variables λ , which are intrinsically random, is to completely explain the correlation between the outcomes of each part. But we can go further and assume that the whole randomness of the system is due to such variables, like if the probabilistic nature of the outcomes is incorporated in the distribution $\pi(\lambda)$ [7]. In this sense, once we know the hidden variable λ , the result of every possible measurement is determined and there is no correlation between the systems.

In this view, the local model works by taking a λ variable from a random source that works according to $\pi(\lambda)$ and using it to compute the response functions p_A, p_B . The integral in Eq. (3.7) means that considering all the λ that provide outcome a for a measurement $\{M_i\}$, we reach $p(a|\{M_i\})$. So we can imagine that exists a random source to which all parts have access providing the variables that Alice and Bob input in their response functions. As we will see, it is part of the strategy of the Gisin-Degorre model that Alice sieves such variables before using them, provoking a bias on the initially uniform distribution of them.

Another remarking difference is that the authors are generally interested in the mean quantity of information exchanged in a process, so that it is natural to consider the expected value of an outcome instead of its probability. This means that now we are going to simulate the expectation $E(A)$, where A is the random outcome of the observable $x \cdot \sigma$, instead of $p(a|x)$ for some fixed a . An immediate consequence is that now our local model will be label-dependent, unlike Werner's.

With the deterministic view of the measurements' results provided by the shared random variables, the local model is now defined by response functions of the form $A(x, \lambda)$ that take values on $\{-1, 1\}$, accordingly to the random variable $\lambda \in \Lambda$, which has distribution $\pi(\lambda)$. In this way, the joint expectation of measurements $x \cdot \sigma$ done by Alice and $y \cdot \sigma$ done by Bob becomes

$$E(AB) = \int_{\Lambda} \pi(\lambda) A(x, \lambda) B(y, \lambda) d\lambda.$$

Now we must also guarantee that the marginals are achieved properly by the response functions, so we must check that

$$E(A) = \int_{\Lambda} A(x, \lambda) \pi(\lambda) d\lambda$$

and

$$E(B) = \int_{\Lambda} B(y, \lambda) \pi(\lambda) d\lambda.$$

6.2 A local model for the EPR experiment correlations

In Section 2.5, we presented the EPR experiment, which proves that if Quantum Theory were local, it would be an incomplete theory. There we used the observables σ_x, σ_z , but any other pair of non-commuting observables could replace them. We now present a version slightly modified of the experiment, where Alice and Bob are allowed to measure the spin along any direction they choose, based on Bohm's simplified version.

Definition 6.1. *Alice and Bob share a qubit pair in the singlet state $|\Psi_{-}\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$. Each of them then receive the description of a projective measurement they have to perform in their respective qubit, which can be represented by unit vectors $|x\rangle$ and $|y\rangle$ in \mathbb{R}^3 . They then obtain measurement outcomes $A, B \in \{1, -1\}$, where 1 corresponds to a spin parallel to the measurement's direction and -1 to spin anti-parallel to this direction.*

According to Quantum Theory, the outcome of Alice's and Bob's measurements, A and B , have the following joint probabilities:

$$p(AB) = \frac{1 - ABx \cdot y}{4},$$

or, equivalently, their joint and marginal expectation values are given by

$$\begin{aligned} E(AB) &= -x \cdot y, \\ E(A) &= 0, \\ E(B) &= 0. \end{aligned}$$

So, in order to simulate locally the EPR experiment, we must achieve these three expected values. Because of the Bell inequality violation (Section 3.1), we know that this cannot be done using only hidden variables, that is, there is no way to define $\lambda, \pi(\lambda), A(x, \lambda)$ and $B(y, \lambda)$ such that

$$E(AB) = \int \pi(\lambda) A(x, \lambda) B(y, \lambda) d\lambda,$$

with $\pi(\lambda)$ independent of x and y .

However, Degorre and co-workers showed that if we allow the distribution of the hidden variables *to depend of Alice's measurement*, then statistics can be reproduced locally. In

order to prove that, we start with two technical lemmas, which will be useful to the next results as well.

Lemma 2. *For any fixed unit vector $x \in \mathbb{R}^3$ we have*

$$\int_{\mathbb{S}^2} (x \cdot \lambda) d\lambda = 0.$$

Proof. Using spherical coordinates and adopting the reference frame where $|x\rangle = (1, 0, 0)$ and $|\lambda\rangle = (1, \theta, \phi)$, we have

$$\int_{\mathbb{S}^2} (x \cdot \lambda) d\lambda = \int_0^\pi \int_0^{2\pi} (x \cdot \lambda) \sin \theta d\phi d\theta.$$

The scalar product of $|x\rangle$ and $|\lambda\rangle$ is the co-sine of the angle between them. With the reference frame we set, by definition this angle is θ , since we put $|a\rangle$ over the z -axis. So we obtain

$$\begin{aligned} \int_0^\pi \int_0^{2\pi} (x \cdot \lambda) \sin \theta d\phi d\theta &= 2\pi \int_0^\pi \cos \theta \sin \theta d\theta \\ &= 2\pi \left[\frac{1}{2} \sin^2 \theta \right]_0^\pi \\ &= 0. \end{aligned}$$

□

Lemma 3. *For any fixed unit vector $x \in \mathbb{R}^3$ we have¹*

$$\int_{\mathbb{S}^2} |x \cdot \lambda| d\lambda = 2\pi.$$

Proof. Under the same assumptions in the proof of Lemma 2, we have

$$\begin{aligned} \int_{\mathbb{S}^2} |x \cdot \lambda| d\lambda &= \int_0^\pi \int_0^{2\pi} |\cos \theta| \sin \theta d\phi d\theta \\ &= 2\pi \int_0^\pi |\cos \theta| \sin \theta d\theta \\ &= 2\pi \left(\int_0^{\frac{\pi}{2}} \cos \theta \sin \theta d\theta - \int_{\frac{\pi}{2}}^\pi \cos \theta \sin \theta d\theta \right) \\ &= 2\pi \left(\left[\frac{1}{2} \sin^2 \theta \right]_0^{\frac{\pi}{2}} - \left[\frac{1}{2} \sin^2 \theta \right]_{\frac{\pi}{2}}^\pi \right) \\ &= 2\pi. \end{aligned}$$

¹This result can be generalized to $\int_{\mathbb{S}^n} |x \cdot \lambda| d\lambda = \frac{2}{n} S_{n-1}$, where S_{n-1} denotes the surface area of \mathbb{S}^{n-1} [26].

□

Theorem 7. *Let $|x\rangle$ and $|y\rangle$ be respectively Alice's and Bob's inputs. If Alice and Bob share a random variable $|\lambda\rangle \in \mathbb{S}^2$ distributed according to a biased distribution with probability density*

$$\pi(\lambda_s|x) = \frac{|x \cdot \lambda_s|}{2\pi} \quad (6.1)$$

then they are able to simulate the correlations of the EPR experiment without any further resource.

Proof. First, notice that π is indeed a probability density function, since $\pi(\lambda) \geq 0$ for all $\lambda \in \mathbb{S}^2$ and Lemma 3 shows that

$$\frac{1}{2\pi} \int_{\mathbb{S}^2} |x \cdot \lambda_s| d\lambda_s = 1.$$

If Alice and Bob set their respective response functions as

$$A(x, \lambda_s) = -\text{sign}(x \cdot \lambda_s), \quad B(y, \lambda_s) = \text{sign}(y \cdot \lambda_s), \quad (6.2)$$

where the sign function $\text{sign}(z)$ is defined as 1 if $z \geq 0$ and -1 if $z < 0$, then the joint expectation value is given by²

$$\begin{aligned} E(AB) &= \int_{\mathbb{S}^2} \pi(\lambda_s|x) A(x, \lambda_s) B(y, \lambda_s) d\lambda_s \\ &= \frac{-1}{2\pi} \int_{\mathbb{S}^2} |x \cdot \lambda_s| \text{sign}(x \cdot \lambda_s) \text{sign}(y \cdot \lambda_s) d\lambda_s \\ &= \frac{-1}{2\pi} \int_{\mathbb{S}^2} (x \cdot \lambda_s) \text{sign}(y \cdot \lambda_s) d\lambda_s \end{aligned}$$

To solve this last integral, we choose a reference frame where $|y\rangle = (0, 0, 1)$, $|x\rangle =$

²Notice that in this case we can also write the sign function $\text{sign}(|x\rangle \cdot |\lambda_s\rangle)$ to be 1 if $\langle x | \lambda_s \rangle > \langle x' | \lambda_s \rangle$ and -1 otherwise, much similar to Bob's response function in the Werner model (Chapter 4).

$(\sin \alpha, 0, \cos \alpha)$ and $|\lambda_s\rangle = (\sqrt{1 - \eta^2} \cos \phi, \sqrt{1 - \eta^2} \sin \phi, \eta)$, which yields

$$\begin{aligned}
\frac{-1}{2\pi} \int_{\mathbb{S}^2} (x \cdot \lambda_s) \text{sign}(y \cdot \lambda_s) d\lambda_s &= \frac{-1}{2\pi} \int_{-1}^1 \int_0^{2\pi} [\sin(\alpha) \cos(\phi) \sqrt{1 - \eta^2} + \cos(\alpha) \eta] \text{sign}(\eta) d\phi d\eta \\
&= \frac{-1}{2\pi} \int_{-1}^1 \left[\sin \alpha \sqrt{1 - \eta^2} \text{sign}(\eta) \int_0^{2\pi} \cos \phi d\phi \right. \\
&\quad \left. + \int_0^{2\pi} \cos(\alpha) \eta \text{sign}(\eta) d\phi \right] d\eta \\
&= \frac{-1}{2\pi} \int_{-1}^1 2\pi \cos(\alpha) \eta \text{sign}(\eta) d\eta \\
&= - \left[\int_{-1}^0 \cos(\alpha) \eta d\eta - \int_0^{-1} \cos(\alpha) \eta d\eta \right] \\
&= -\cos \alpha \left(\left[\frac{\eta^2}{2} \right]_{-1}^0 - \left[\frac{\eta^2}{2} \right]_0^{-1} \right) \\
&= -\cos \alpha \\
&= -x \cdot y.
\end{aligned}$$

Since $E(A) = 0$ is almost exactly the content of Lemma 2, to finish the proof we only have to establish $E(B) = 0$. Let's consider the half-spheres $\mathbb{S}_+, \mathbb{S}_-$ with respect to x and $\lambda_- \in \mathbb{S}_-, \lambda_+ \in \mathbb{S}_+$. We have

$$\begin{aligned}
E(B) &= \int_{\Lambda} B(y, \lambda_s) \pi(\lambda_s | x) d\lambda_s \\
&= \frac{1}{2\pi} \int_{\Lambda} |x \cdot \lambda_s| \text{sign}(y \cdot \lambda_s) d\lambda_s \\
&= \frac{1}{2\pi} \left[\int_{\mathbb{S}_+} (x \cdot \lambda_+) \text{sign}(y \cdot \lambda_+) d\lambda_+ - \int_{\mathbb{S}_-} (x \cdot \lambda_-) \text{sign}(y \cdot \lambda_-) d\lambda_- \right].
\end{aligned}$$

Substituting $\lambda_- = -\lambda_+$, we obtain

$$E(B) = \frac{1}{2\pi} \left[\int_{\mathbb{S}_+} (x \cdot \lambda_+) \text{sign}(y \cdot \lambda_+) d\lambda_+ - \int_{\mathbb{S}_+} (x \cdot \lambda_+) \text{sign}(y \cdot \lambda_+) d\lambda_+ \right] = 0.$$

□

The above theorem shows that with a biased distribution of the hidden variables, Alice and Bob are able to do simulate the correlations of the EPR experiment locally. Nevertheless, we can consider the situation where we start with a uniform distribution of the λ , just like in Werner's model, and then generate the bias by a sampling of the hidden variables performed by Alice.

The process to carry on this task can be divided in two steps. First Alice samples from the uniformly distributed random variables λ the biased variable λ_s using her knowledge of a , with the bias given by Eq. (6.1). In other words, Alice tests the hidden variable provided by the source (initially unbiased) and accepts it or discards it, in such a way to produce the probability distribution (6.1). The second step is the communication complexity problem of communicate Bob about which of the hidden variables was accepted. It is to accomplish this second step that we will need additional resources to those allowed by a local model.

6.3 The sampling

To performing the sampling, Alice uses the “choice method”³, performed in the following way. Once Alice receives λ_0 and λ_1 from the random source, she calculates $|a \cdot \lambda_0|$ and $|a \cdot \lambda_1|$. The variable which provides the higher value is accepted as λ_s .

Protocol 1. (*Choice method*)

1. Alice picks $\lambda_0 \sim \text{Unif}(\mathbb{S}^2)$
2. Alice picks $\lambda_1 \sim \text{Unif}(\mathbb{S}^2)$
3. If $|x \cdot \lambda_0| > |x \cdot \lambda_1|$, then she accepts λ_0 and sets $\lambda_s = \lambda_0$; otherwise she accepts λ_1 and sets $\lambda_s = \lambda_1$.

Theorem 8. *If Alice performs the choice method protocol, then $\lambda_s \sim |x \cdot \lambda_s|/2\pi$ and $p(\lambda_s = \lambda_1) = p(\lambda_s = \lambda_2) = 1/2$.*

Proof. A uniform distribution of λ in \mathbb{S}^2 implies a uniform distribution of $|x \cdot \lambda|$ in the interval $[0, 1]$, therefore each λ_i is accepted with probability $|x \cdot \lambda_i|$. Since the probability density function of the uniform distribution on \mathbb{S}^2 is $\frac{1}{2\pi}$, after the protocol we have $\lambda_s \sim |x \cdot \lambda_s|/2\pi$. \square

6.4 The communication

Once Alice uses the choice method, Theorem 8 guarantees that the correct probability distribution (6.1) occurs. In order to apply Theorem 7 to conclude that the EPR experiment correlations are reproduced, all we need now is to make sure that Bob also knows which is the sampled hidden variable λ_s , so that he can evaluate the appropriate response function $B(y, \lambda_s) = \text{sign}(y \cdot \lambda_s)$.

³In [9], they initially present the “rejection method”, where Alice rejects λ_k if $|a \cdot \lambda_k| < u_k$, where $u_k \sim \text{Unif}[0, 1]$. The big drawback is that Alice could reject an arbitrarily large number of samples before accepting one, while in the choice method only one round is enough to perform the sampling. The equivalence of the bias induced by both methods can be seen by noticing that $|a \cdot \lambda| \sim \text{Unif}[0, 1]$

Considering only what a standard local hidden variables model allows - that is, shared randomness - the Bell inequality violation guarantees that there is no way to tell Bob about the result of the sampling, since Alice does everything locally. However, in the case where we allow classical communication between both parts, the problem resumes to be trivial. Actually, more: the trivial answer is even optimal. Once the sampling is done, all that Alice has to do is to send one (classical) bit $s \in \{0, 1\}$ to Bob saying whether $\lambda_s = \lambda_0$ or $\lambda_s = \lambda_1$.

The problem of all classical communication to shared randomness with the objective of reproduce quantum correlations has been studied since 1992, when Tim Maudlin showed that, in average, a finite amount of bits would suffice [27]. In 2003, Ben Toner and Dave Bacon got the most effective optimization of it, presenting he first local model that, together with one bit, simulate successfully the correlations of the singlet [28].

6.5 A local model for the Werner state

With this framework in hands, only a simple observation is enough for us to show that we have a local model (without any further resources) for the $\mathbb{C}^2 \otimes \mathbb{C}^2$ Werner state

$$W_{2 \times 2} = \alpha |\Psi_{-}\rangle \langle \Psi_{-}| + (1 - \alpha) \frac{I_{4 \times 4}}{4} \quad (6.3)$$

with $\alpha = 1/2$.

By linearity, we can see that this state generates the joint expectation value $E(AB) = -(x \cdot y)/2$.

Giving up the possibility to use classical communication, we come back to the weaker assumption that Alice and Bob have only shared randomness and, therefore, to the situation where Bob is not aware of Alice's sampling, so he cannot distinguish λ_s from λ_0 (or λ_1). The point is that if Bob always evaluate his respective response function assuming $\lambda_s = \lambda_0$, since $p(\lambda_s = \lambda_0) = p(\lambda_s = \lambda_1) = 1/2$ on average, he will be right half of the time and thus the correlations will match those of the singlet. In the other half, the independent response functions will generate the probabilities achieved by maximal random noise, that is, by the state $I_{4 \times 4}/4$.

Thus we are proposing to use the following protocol.

Protocol 2. (*Local model for the state $W_{2 \times 2}$*)

1. Alice and Bob share a pair of variables $\lambda_0, \lambda_1 \sim \text{Unif}(\mathbb{S}^2)$.
2. Alice performs the choice method applied to λ_0, λ_1 and outputs $A = -\text{sign}(x \cdot \lambda_s)$.
3. Bob outputs $B = \text{sign}(y \cdot \lambda_0)$.

We now proceed to show that the local model indeed reproduces the correlations of

state $W_{2 \times 2}$.

Lemma 4. *Alice's response function (6.2) can be rewritten as*

$$A(x, \lambda_s) = -\text{sign}(x \cdot (\lambda_0 + \lambda_1)).$$

Proof. Notice that the sign function allows us to write

$$-\text{sign}(x \cdot \lambda) = (-1)^{\chi_{S_x}(\lambda)},$$

where χ_S is the characteristic function of the set S (i.e., $\chi_S(s)$ equals 1 if $s \in S$ and 0 otherwise) and S_x denotes the semi-sphere $\{\lambda \in \mathbb{S}^2; x \cdot \lambda \geq 0\}$.

From the definition of λ_s in the choice method protocol, we conclude that

$$\chi_{S_x}(\lambda_0 + \lambda_1) = \chi_{S_x}(\lambda_s),$$

which means that $-\text{sign}(x \cdot \lambda_s) = -\text{sign}(x \cdot (\lambda_0 + \lambda_1))$. □

Theorem 9. *There exists a local model for the Werner state*

$$W_{2 \times 2} = \frac{1}{2} |\Psi_{-}\rangle \langle \Psi_{-}| + \frac{1}{2} \frac{I_{4 \times 4}}{4}.$$

Proof. As already stated, the local model consists of Alice performing the choice method and Bob always assuming that $\lambda_s = \lambda_0$. Using Lemma 4, Lemma 2 and adjusting the differential, we obtain

$$\begin{aligned} E(AB) &= -\frac{1}{8\pi^2} \int_{\mathbb{S}^2} d\lambda_0 \int_{\mathbb{S}^2} d\lambda_1 (x \cdot (\lambda_0 + \lambda_1)) \text{sign}(y \cdot \lambda_0) \\ &= -\frac{1}{8\pi^2} 2\pi(x \cdot y) \int_{\mathbb{S}^2} d\lambda_1 \\ &= -\frac{x \cdot y}{2}. \end{aligned}$$

As we also have $E(A) = E(B) = 0$, the model reproduces the correlations exhibited by $W_{2 \times 2}$. □

Chapter 7

Genuine hidden nonlocality

So far, this is the road we have been following: Werner constructed a local model for projective measurements for the state W_{local} ; Barrett constructed a local model for POVMs for the state \hat{W} , similar to W_{local} but not the same; and Popescu showed that states W_{local} of dimension $d \geq 5$ have hidden nonlocality revealed by a proper sequence of measurements. Since it is not known if W_{local} is local for POVMs or not, these facts are not enough to assure that a sequence of measurements is really necessary to activate its nonlocality. If we adopt the natural hierarchy where a single POVM has a lower “cost” than a couple of projective measurements, then is reasonable to say that Popescu’s activation protocol still left room for optimization. Are two measurements indeed necessary?

On the other hand, it is not known if Barrett’s state \hat{W} does not also present hidden nonlocality; so far no-one was able to find a sequence of measurements to display it, but this does not mean that it is impossible to exist one. The natural question raised here is: is there an entangled state, the nonlocality of which can be observed only if sequential measurements are used?

In the paper “*Genuine hidden quantum nonlocality*” [10], which is the subject of this chapter, the term “Genuine” holds in the above sense. Brunner and co-workers presented a state which counts with a local model for POVMs, which nevertheless can be shown to violate the CHSH inequality after a sequence of judiciously chosen local measurements are performed. It was the first and, until present date, the only example of a genuine necessity of more than one measurement.

7.1 A local model for dichotomic projective measurements

We start by the construction of a local model for dichotomic projective measurements for a state of $\mathbb{C}^2 \otimes \mathbb{C}^2$. More specifically, we consider the class of states

$$\begin{aligned}\rho_G(q) &= q |\Psi_-\rangle \langle \Psi_-| + (1-q) |0\rangle \langle 0| \otimes \frac{I_{2 \times 2}}{2} \\ &= q |\Psi_-\rangle \langle \Psi_-| + \frac{1-q}{2} (|00\rangle \langle 00| + |01\rangle \langle 01|),\end{aligned}\tag{7.1}$$

recalling that $|\Psi_-\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$ is the singlet state. We will see that state (7.1) admits a local model mentioned above when $q \leq 1/2$. Nevertheless, we can use the flip operator as witness, as we did with the Werner states in Theorem 6, to prove that the state is entangled for¹ $q > 1/3$:

$$\begin{aligned}\text{tr}(V \rho_G) &= q \text{tr}(V |\Psi_-\rangle \langle \Psi_-|) + \frac{1-q}{2} (\text{tr}(V |00\rangle \langle 00|) + \text{tr}(V |01\rangle \langle 01|)) \\ &= \frac{q}{2} (\text{tr}(V |01\rangle \langle 01|) + \text{tr}(V |10\rangle \langle 10|) - \text{tr}(V |10\rangle \langle 01|) - \text{tr}(V |01\rangle \langle 10|)) + \frac{1-q}{2} \\ &= -q + \frac{1-q}{2},\end{aligned}$$

which is negative if and only if $q > 1/3$.

As in Chapter 6, the statistics we wish to simulate for ρ_G are $E(A)$, $E(B)$ and $E(AB)$. Notice that, by linearity,

$$E_{\rho_G}(A) = q E_{|\Psi_-\rangle \langle \Psi_-|}(A) + (1-q) E_{|0\rangle \langle 0| \otimes \frac{I}{2}}(A)$$

and analogously for $E(B)$ and $E(AB)$, where the indices say which state should be considered for each expectation. In Chapter 6, we saw that $E_{|\Psi_-\rangle \langle \Psi_-|}(A) = E_{|\Psi_-\rangle \langle \Psi_-|}(B) = 0$, thus $E(A) = (1-q) E_{|0\rangle \langle 0|}(A)$ and $E(B) = (1-q) E_{\frac{I}{2}}(B)$. Therefore, we have

$$\begin{aligned}E(A) &= (1-q) \left[\frac{1 + \langle 0 | x \cdot \sigma | 0 \rangle}{2} - \frac{1 - \langle 0 | x \cdot \sigma | 0 \rangle}{2} \right] \\ &= (1-q) \langle 0 | x \cdot \sigma | 0 \rangle \\ &= (1-q) x_z\end{aligned}$$

¹Actually, the state is entangled for all $q > 0$, as can be seen using the Peres-Horodecki Criterion [13].

(where $|x\rangle = (x_x, x_y, x_z)$) and

$$\begin{aligned} E(B) &= (1-q) \left[\text{tr} \left(\frac{I}{2} \frac{I + y \cdot \sigma}{2} \right) - \text{tr} \left(\frac{I}{2} \frac{I - y \cdot \sigma}{2} \right) \right] \\ &= \frac{(1-q)}{2} \text{tr}(y \cdot \sigma) \\ &= 0. \end{aligned}$$

For $E(AB)$, since $|0\rangle\langle 0| \otimes \frac{I}{2}$ is a product state (thus the outcomes of any local measurements are uncorrelated), we have

$$\begin{aligned} E(AB) &= q E_{|\Psi_-\rangle\langle\Psi_-|}(AB) + (1-q) E_{|0\rangle\langle 0| \otimes \frac{I}{2}}(AB) \\ &= q E_{|\Psi_-\rangle\langle\Psi_-|}(AB) + (1-q) E_{|0\rangle\langle 0|}(A) E_{\frac{I}{2}}(B) \\ &= q E_{|\Psi_-\rangle\langle\Psi_-|}(AB) \\ &= -q(x \cdot y), \end{aligned}$$

where the last equality was shown in Chapter 6.

The performance of a dichotomic projective measurement is precisely what is done in the EPR experiment (see Section 2.5), the only difference is the state being shared: now we are using the modified singlet state $\rho_G(q)$, instead of the singlet $\rho_G(1)$ itself. Alice (and Bob) will receive as input a vector $|x\rangle$ (and $|y\rangle$), and should simulate the statistics of measuring the qubit observables $x \cdot \sigma$ and $y \cdot \sigma$, with possible outcomes $A, B \in \{-1, 1\}$. The following protocol is strongly inspired in Protocol 2.

Protocol 3 (Simulation of $\rho_G(1/2)$). *1. Alice and Bob share a tridimensional unit vector $|\lambda\rangle$, uniformly distributed on the sphere.*

2. Upon receiving x , Alice tests λ . With probability $|x \cdot \lambda|$, she accepts λ and outputs $A = -\text{sign}(x \cdot \lambda)$; otherwise, she outputs $A = \pm 1$ with probability $(1 \pm \langle 0 | x \cdot \sigma | 0 \rangle)/2$.

3. Bob outputs $B = \text{sign}(y \cdot \lambda)$.

Theorem 10. *There exists a local model that simulates the correlations exhibited by state $\rho_G(q)$ upon the measurement of a dichotomic projective measurement, for any $q \in [0, 1/2]$.*

Proof. Let $x \cdot \sigma, y \cdot \sigma$ be the observables measured by Alice and Bob, respectively. After Protocol 3 is performed, as shown in Theorem 8, if Alice accepts λ , which occurs on average with probability $1/2$ (independently of x), λ is distributed according to the density $\pi(\lambda) = |x \cdot \lambda|/2\pi$. In this case, since Bob outputs $B = \text{sign}(y \cdot \lambda)$, Theorem 7 says that

the correlation between Alice's and Bob's outcomes is

$$\begin{aligned} E(AB) &= \frac{-1}{2\pi} \int_{\mathbb{S}_2} |x \cdot \lambda| \operatorname{sign}(x \cdot \lambda) \operatorname{sign}(y \cdot \lambda) \\ &= -x \cdot y. \end{aligned}$$

As the marginal expectations are uniform, i.e., $E(A) = E(B) = 0$, we recover the singlet correlations.

If Alice rejects λ , she simulates the statistics of state $|0\rangle$, while Bob's outcome is uncorrelated and uniformly distributed. Hence the model reproduces exactly the statistics of state (7.1) for $q = 1/2$.

For $0 < q < 1/2$, it is sufficient to observe that for $q = 0$ the state is clearly local and that the set of all local states is convex. But since we haven't discussed such topics in this text (we recommend for the interested reader [7], or [29] for a gentler introduction), we will explicitly show such convexity for this class of states, presenting a simple protocol of simulation.

Fix $p \in [0, 1]$. The following protocol simulates the state $\check{\rho}(p) = p\rho_G(1/2) + (1 - p)|0\rangle\langle 0| \otimes \frac{I}{2}$. (Notice that $|0\rangle\langle 0| \otimes \frac{I}{2} = \rho_G(0)$.)

Protocol 4 (Simulation of $\check{\rho}(p)$). *1. Alice and Bob share a real number r uniformly distributed on the interval $[0, 1]$ and a tridimensional unit vector $|\lambda\rangle$, uniformly distributed on the sphere.*

2. With probability x , Alice accepts r and executes Protocol 3; otherwise she outputs $A = \pm 1$ with probability $(1 \pm \langle 0|x \cdot \sigma|0\rangle)/2$.

3. Bob outputs $B = \operatorname{sign}(y \cdot \lambda)$.

Now, one just have to check that

$$\begin{aligned} \check{\rho}(p) &= p\rho_G(1/2) + (1 - p)|0\rangle\langle 0| \frac{I}{2} \\ &= \frac{p}{2} |\psi_{-}\rangle\langle\psi_{-}| + \left(\frac{p}{2} + (1 - p)\right) |0\rangle\langle 0| \frac{I}{2} \\ &= \frac{p}{2} |\psi_{-}\rangle\langle\psi_{-}| + \left(1 - \frac{p}{2}\right) |0\rangle\langle 0| \frac{I}{2} \\ &= \rho_G(p/2). \end{aligned}$$

Therefore, choosing the proper p on $[0, 1]$, with Protocol 4 we can simulate $\rho_G(q)$ for any $q \in [0, 1/2]$. \square

7.2 Revealing nonlocality

Similar to what we have done in Chapter 5, we are now going to show that after local filtering, the state ρ_G violates the CHSH inequality (see Section 3.1).

More specifically, we are going to apply filters of the form

$$F_A = \epsilon |0\rangle \langle 0| + |1\rangle \langle 1|, \quad F_B = \delta |0\rangle \langle 0| + |1\rangle \langle 1| \quad (7.2)$$

with $\delta = \epsilon/\sqrt{q}$ to ρ_G . In another words, we will perform the measurements $\{F_A, I - F_A\}$ in Alice's side and $\{F_B, I - F_B\}$ in Bob's, and discard the protocol in case that the outcome is referent to operator $I - F_A$ or $I - F_B$. This means that, after filtering, the resulting (unnormalized) state is

$$\begin{aligned} \tilde{\rho}_G &= F_A \otimes F_B \rho_G F_A^\dagger \otimes F_B^\dagger \\ &= q F_A \otimes F_B |\Psi_-\rangle \langle \Psi_-| F_A \otimes F_B + (1-q)(F_A |0\rangle \langle 0| F_A) \otimes (F_B \frac{I}{2} F_B) \\ &= \frac{q}{2} \left[\epsilon^2 |01\rangle \langle 01| + \frac{\epsilon^2}{q} |10\rangle \langle 10| - \frac{\epsilon^2}{\sqrt{q}} (|01\rangle \langle 10| + |10\rangle \langle 01|) \right] \\ &\quad + \frac{1-q}{2} \left[\frac{\epsilon^4}{q} |00\rangle \langle 00| + \epsilon^2 |01\rangle \langle 01| \right] \\ &= \epsilon^2 \left[\frac{1}{2} |10\rangle \langle 10| - \frac{\sqrt{q}}{2} (|01\rangle \langle 10| + |10\rangle \langle 01|) + \frac{1}{2} |01\rangle \langle 01| \right] + O(\epsilon^4) \\ &\simeq \frac{1}{2} |10\rangle \langle 10| + \frac{1}{2} |01\rangle \langle 01| - \frac{\sqrt{q}}{2} (|01\rangle \langle 10| + |10\rangle \langle 01|) + O(\epsilon^2) \end{aligned}$$

By adding and subtracting to both sides $\sqrt{q}(|01\rangle \langle 01| + |10\rangle \langle 10|)/2$, we achieve

$$\tilde{\rho}_G \simeq \sqrt{q} |\Psi_-\rangle \langle \Psi_-| + (1 - \sqrt{q}) \frac{|01\rangle \langle 01| + |10\rangle \langle 10|}{2} + O(\epsilon^2).$$

According to the Horodecki criterion (see Section 3.2), we can calculate that $\tilde{\rho}_G$ violates CHSH up to $2\sqrt{1+q}$ (for $\epsilon \rightarrow 0$). Thus, state $\rho_G(q)$ is local for all $0 \leq q \leq 1/2$ and exhibits hidden nonlocality for projective measurements for all $q > 0$. In the unique other know example of hidden nonlocality (Chapter 5), the local dimension was $d \geq 5$, making this case (with local dimension $d = 2$) the simplest example of the phenomenon.

7.3 A local model for POVMs

Nevertheless, our main goal was not achieved yet. We cannot guarantee that such nonlocality is genuine, in the sense expressed in the beginning of the chapter: the local model constructed accounts only for projective measurements. In principle, a Bell violation can

be obtained using POVMs.

However, we now proceed to present a protocol for the construction of a state which actually admits a local model for POVMs, based on another one which is local for projective measurements. To be more precise, once the local model for projective measurements is done for the initial state, we will show that a repeated usage of it provides the simulation of the POVMs for the second one. In this sense, it represents an optimization of utilization of the hidden variables and the response functions of the initial model.

Admitting that a given state ρ_0 of local dimension d is local for projective dichotomic measurements, consider the state

$$\rho' = \frac{1}{d^2}[\rho_0 + (d-1)(\rho_A \otimes \sigma_B + \sigma_A \otimes \rho_B) + (d-1)^2 \sigma_A \otimes \sigma_B], \quad (7.3)$$

where σ_A, σ_B are arbitrary d -dimensional states and $\rho_A = \text{tr}_B(\rho_0), \rho_B = \text{tr}_A(\rho_0)$. (ρ' is indeed a state since is convex combination of states.)

Suppose that Alice and Bob receives as input the POVMs $M = \{M_a\}$ and $N = \{N_b\}$, respectively. Then the expected value $p(ab)$ which we are interested to reproduce locally is given by

$$\begin{aligned} \text{tr}(M_a \otimes N_b \rho') &= \frac{1}{d^2}[\text{tr}(M_a \otimes N_b \rho_0) + (d-1) \text{tr}(M_a \rho_A) \text{tr}(N_b \sigma_B) \\ &\quad + (d-1) \text{tr}(M_a \sigma_A) \text{tr}(N_b \rho_B) + (d-1)^2 \text{tr}(M_a \sigma_A) \text{tr}(N_b \sigma_B)]. \end{aligned} \quad (7.4)$$

Following Proposition 4.4.1, we can assume that the elements of both POVMs are proportional to one-dimensional projectors, i.e., $M_a = \alpha_a P_a$ and $N_b = \beta_b Q_b$, with $\alpha_a, \beta_b \geq 0$. Note that by normalization of the POVM,

$$I = \sum_a M_a = \sum_a \alpha_a P_a,$$

which implies

$$d = \text{tr}(I) = \sum_a \alpha_a \text{tr}(P_a) = \sum_a \alpha_a.$$

Similarly, we find $\sum_b \beta_b = d$.

We will show that ρ' is local for POVMs through the following protocol. The protocol is written for Alice, but Bob follows the same procedure.

- Protocol 5.**
1. Alice chooses projector P_a with probability α_a/d (notice that $\sum_a \alpha_a/d = 1$).
 2. She simulates the dichotomic projective measurement $\{P_a, I - P_a\}$ on state ρ_0 .
 3. If the output of the simulation corresponds to P_a , she outputs a .

4. Otherwise, she outputs (any) a with probability $\text{tr}(M_a \sigma_A)$.

Theorem 11. *If there is a local model for dichotomic measurements over ρ_0 , then there is a local model that simulates the correlations exhibited by state ρ' given in Eq. (7.3) upon the measurement of local POVMs.*

Proof. Suppose Alice and Bob are able to simulate locally the correlations of projective dichotomic measurements for the state ρ_0 . Let the Protocol 5 be performed and fix a, b of the set of possible outcomes. Our goal is to show that $p(a, b) = \text{tr}(M_a \otimes N_b \rho')$, where the left side probability is calculated according to the protocol.

First, notice that the probability that Alice (and the same holds for Bob) outputs in step 3 (any outcome) is

$$\sum_a \frac{\alpha_a}{d} \text{tr}(P_a \rho_A) = \frac{1}{d} \sum_a \text{tr}(M_a \rho_A) = \frac{1}{d}.$$

Since each part can output in step (3) or in step (4), four possibilities to obtain outputs a and b arise.

- Both Alice and Bob output in step 3, which occurs with probability

$$\frac{\alpha_a}{d} \frac{\beta_b}{d} \text{tr}(P_a \otimes Q_b \rho_0) = \frac{1}{d^2} \text{tr}(M_a \otimes M_b \rho_0);$$

- Alice outputs in step 3 and Bob in step 4. Since Bob outputs in step 4 with probability $(d-1)/d$, this occurs with probability

$$\frac{\alpha_a}{d} \text{tr}(P_a \rho_A) \frac{d-1}{d} \text{tr}(N_b \sigma_B) = \frac{d-1}{d^2} \text{tr}(M_a \rho_A) \text{tr}(N_b \sigma_B);$$

- Bob outputs in step 3 and Alice in step 4, which occurs with probability

$$\frac{d-1}{d^2} \text{tr}(N_b \rho_B) \text{tr}(M_a \sigma_A);$$

- Both output in step 4, which occurs with probability

$$\frac{(d-1)^2}{d^2} \text{tr}(M_a \sigma_A) \text{tr}(N_b \sigma_B).$$

Altogether, we find that $p(a, b)$ matches accurately Eq. (7.4). □

7.4 Revealing genuine nonlocality

Theorem 11 says that the same local model that reproduces the correlations of dichotomic projective measurements for ρ_0 can reproduce the correlations of a POVM applied to ρ' , if the right protocol is executed. We are now going to use this result to compile everything we saw in the last sections and construct a local state for POVM which violates the CHSH inequality after filtering, proving that a sequence of measurements is indeed necessary in certain cases.

In Eq. (7.3), if we set the local dimension to be $d = 2$ and $\rho_0 = \rho_G(q)$ (given in Eq. (7.1)), which is local for projective measurements for $q \leq 1/2$, we will obtain the reduced states

$$\begin{aligned}\rho_A(q) &= q \frac{I_{2 \times 2}}{2} + (1 - q) |0\rangle \langle 0|, \\ \rho_B(q) &= q \frac{I_{2 \times 2}}{2} + (1 - q) \frac{I_{2 \times 2}}{2} = \frac{I_{2 \times 2}}{2}.\end{aligned}$$

Setting also $\sigma_A = \sigma_B = |0\rangle \langle 0|$, we define the state

$$\begin{aligned}\rho'_G(q) &= \frac{1}{4} \left[q |\Psi_-\rangle \langle \Psi_-| + (1 - q) |0\rangle \langle 0| \otimes \frac{I_{2 \times 2}}{2} + q \frac{I_{2 \times 2}}{2} \otimes |0\rangle \langle 0| + (1 - q) |00\rangle \langle 00| \right. \\ &\quad \left. + |0\rangle \langle 0| \otimes \frac{I_{2 \times 2}}{2} + |00\rangle \langle 00| \right] \\ &= \frac{1}{4} \left[q |\Psi_-\rangle \langle \Psi_-| + (2 - q) |0\rangle \langle 0| \otimes \frac{I_{2 \times 2}}{2} + q \frac{I_{2 \times 2}}{2} \otimes |0\rangle \langle 0| + (2 - q) |00\rangle \langle 00| \right].\end{aligned}$$

We know that, by construction, $\rho'_G(q)$ with $q \leq 1/2$ is local, concerning to general measurements. We are now going to show that, despite of that, it violates the CHSH inequality after filtering, therefore exhibiting genuine hidden nonlocality.

Applying filters of the form (7.2) with $\delta = \epsilon/\sqrt{q}$ to state ρ'_G , we obtain the unnormalized state

$$\begin{aligned}\tilde{\rho}'_G &= F_A \otimes F_B \rho'_G F_A \otimes F_B \\ &= \frac{1}{4} \left[q F_A \otimes F_B |\Psi_-\rangle \langle \Psi_-| F_A \otimes F_B + (2 - q) (F_A |0\rangle \langle 0| F_A) \otimes (F_B \frac{I}{2} F_B) \right. \\ &\quad \left. + q (F_A \frac{I}{2} F_A) \otimes (F_B |0\rangle \langle 0| F_B) + (2 - q) (F_A |0\rangle \langle 0| F_A) \otimes (F_B |0\rangle \langle 0| F_B) \right]\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{4} \left[\frac{q}{2} [\epsilon^2 |01\rangle \langle 01| + \frac{\epsilon^2}{q} |10\rangle \langle 10| - \frac{\epsilon^2}{\sqrt{q}} (|01\rangle \langle 10| + |10\rangle \langle 01|)] \right. \\
&\quad \left. + \frac{2-q}{2} \left(\frac{\epsilon^4}{q} |00\rangle \langle 00| + \epsilon^2 |01\rangle \langle 01| \right) \right. \\
&\quad \left. + \frac{q}{2} \left(\frac{\epsilon^4}{q} |00\rangle \langle 00| + \frac{\epsilon^2}{q} |10\rangle \langle 10| + \frac{2-q}{2} \frac{\epsilon^4}{q} |00\rangle \langle 00| \right) \right] \\
&= \frac{\epsilon^2}{4} \left[\frac{q}{2} |01\rangle \langle 01| + \frac{1}{2} |10\rangle \langle 10| - \frac{\sqrt{q}}{2} (|01\rangle \langle 10| + |10\rangle \langle 01|) + |01\rangle \langle 01| \right. \\
&\quad \left. - \frac{q}{2} |01\rangle \langle 01| + \frac{1}{2} |10\rangle \langle 10| \right] + O(\epsilon^4) \\
&\simeq \frac{1}{2} \left[|01\rangle \langle 01| + |10\rangle \langle 10| - \frac{\sqrt{q}}{2} (|01\rangle \langle 10| + |10\rangle \langle 01|) \right] + O(\epsilon^2) \\
&= \frac{|01\rangle \langle 01| + |10\rangle \langle 10|}{2} - \frac{\sqrt{q}}{2} \frac{|01\rangle \langle 10| + |10\rangle \langle 01|}{2} + O(\epsilon^2) \\
&= \frac{\sqrt{q}}{2} |\Psi_-\rangle \langle \Psi_-| + \left(1 - \frac{\sqrt{q}}{2} \right) \frac{|01\rangle \langle 01| + |10\rangle \langle 10|}{2} + O(\epsilon^2).
\end{aligned}$$

The resulting state $\tilde{\rho}'_G$ violates the CHSH inequality up to $2\sqrt{1+q/4}$ (for $\epsilon \rightarrow 0$) according to the Horodecki criterion. Hence, sequential measurements are necessary to reveal the nonlocality of ρ'_G , which therefore exhibits genuine hidden nonlocality.

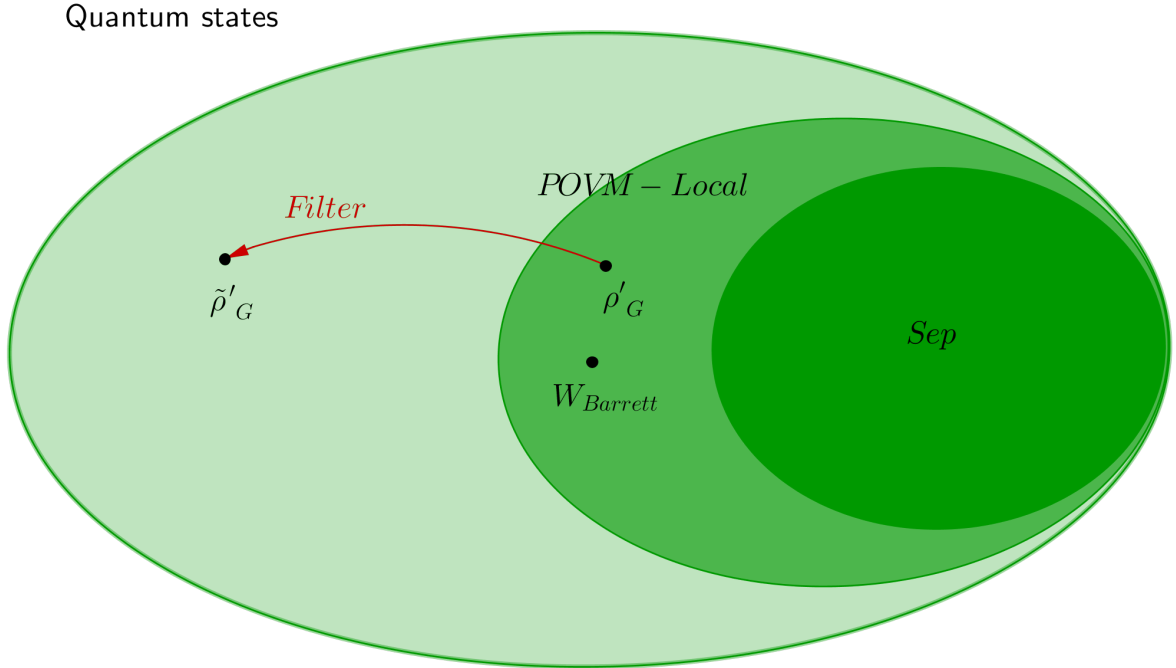


Figure 7.1: The successful application of a local filtering (represented by the arrow in red) on the POVM-local state ρ'_G results in a state $\tilde{\rho}'_G$ which violates CHSH, and thus does not belong to the set of local states.

7.5 Genuine and maximal

Until now, throughout this chapter we have dealt with quantum states of $\mathbb{C}^2 \otimes \mathbb{C}^2$. We conclude the chapter by saying that in [10] the authors also present an even more extremal case of this phenomenon, occurring with a state of $\mathbb{C}^3 \otimes \mathbb{C}^2$, a qutrit-qubit state. Using the presented machinery, they showed that a state with genuine hidden nonlocality can violate maximally the CHSH inequality.

The state

$$\rho_E = q |\Psi_-\rangle \langle \Psi_-| + (1 - q) |2\rangle \langle 2| \otimes \frac{I_2}{2}$$

can be shown to be local for dichotomic projective measurements (a protocol similar to Protocol 3 should be performed). Then, applying Protocol 5 to ρ_E and taking $\sigma_A = \sigma_B = |2\rangle \langle 2|$ we can simulate POVMs on the state

$$\tilde{\rho}_E = \frac{1}{9} \left[q |\Psi_-\rangle \langle \Psi_-| + (3 - q) |2\rangle \langle 2| \otimes \frac{I_2}{2} + 2q \frac{I_2}{2} |2\rangle \langle 2| + (6 - 2q) |22\rangle \langle 22| \right], \quad (7.6)$$

constructed via Eq. (7.3).

To reveal the nonlocality of the above state, we apply filters of the form $F_A = F_B = |0\rangle \langle 0| + |1\rangle \langle 1|$. After successful filtering, the pure singlet state $|\Psi_-\rangle \langle \Psi_-|$ is obtained, which we know to violate maximally CHSH. Hence, state (7.6) has genuine and maximal hidden nonlocality.

Conclusion

In this master's thesis, our focus was to present some examples of local models and hidden nonlocality, providing the calculations underneath them and emphasizing the motivations as much as possible. In particular, the most seminal work on the subject, Werner's local model, was originally presented in a very knotty way. Perhaps a more detailed study of it makes easier to introduce new researchers to the topic; this work maybe is a first step towards that direction.

Naturally, several works about the subject do not appear here, and others are just shortly cited. Specifically in the case of local models, this is a bit frustrating, since the universe of local models is not very large. Barrett's model [25], for example, deserves more space, given its importance. Another important work is Ref. [30], in which it is studied the robustness of nonlocality to noise and it is presented a local model for isotropic states, which inspired Hirsch *et al.* [10]. For the tripartite case, Ref. [31] presents a very interesting local model for projective measurements. A very complete and up-to-date review on the whole local models topic can be found in Ref. [33]. About activation of nonlocality, we left a whole branch untouched, where are considered multiple copies of a state and quantum networks, in order to culminate in a Bell inequality violation [7, 29].

What becomes clear is that this is a research area with a lot of potential, where representative examples are welcome and, with rare exceptions, general results still are only conjectured. For instance, is it possible to create a local model with hidden variables other than unit vectors? Is there any projective-local state which cannot simulate a POVM, meaning that POVMs do offer advantage in order to detect the nonlocality of quantum states? Is it possible to construct a local model for POVMs, in the multipartite case? Is there a quantum entangled state completely local, that is, that does not violate any Bell inequality, even in these more general scenarios of sequences of measurements and multiple copies? Or are entanglement and nonlocality, after all, one and the same thing, in this broader sense?² There is work to be done.

²Related to this last question, there is the Peres Conjecture, which claims that states with bound entanglement are completely local. Recently, the conjecture was disproved [32].

Bibliography

- [1] R. F. Werner, Phys. Rev. A **40**, 4277 (1989).
- [2] J. S. Bell, Physics **1**, 195 (1964).
- [3] S. Popescu, Phys. Rev. Lett. **74**, 2619 (1995).
- [4] A. Einstein, B. Podolsky and N. Rosen, Phys. Rev. **47**, 777 (1935).
- [5] J. F. Clauser, M. A. Horne, A. Shimony and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).
- [6] R. Horodecki, P. Horodecki, M. Horodecki, Phys. Lett. A **200**, 340 (1995).
- [7] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani and S. Wehner, Rev. Mod. Phys. **86**, 419 (2014).
- [8] N. Gisin, B. Gisin, Phys. Lett. A **260**, 323 (1999).
- [9] J. Degorre, J. Roland, S. Laplante, Phys. Rev. A **72**, 062314 (2005).
- [10] F. Hirsch, M. T. Quintino, J. Bowles and N. Brunner, Phys. Rev. Lett. **111**, 160402 (2013).
- [11] E. L. Lima, Álgebra Linear, IMPA (2012).
- [12] B. James, Probabilidade: um curso em nível intermediário, IMPA (2010).
- [13] M. A. Nielsen and I. L Chuang, Quantum Computation and Quantum Information, Bookman (2005).
- [14] B. Amaral, A. T. Baraviera and M. O. Terra Cunha, Mecânica Quântica para matemáticos em formação, IMPA (2011).
- [15] M. O. Terra Cunha, Noções de Informação Quântica, IMPA (2007).
- [16] A. M. Gleason, J. Ration. Mech. Anal. **6**, 885 (1957).

- [17] J. S. Bell, Rev. Mod. Phys. **38** 447-452 (1966).
- [18] M. Araújo, master's thesis, available at arXiv:1208.6283 (2012).
- [19] D. Bohm, Quantum Theory, Dover (1951).
- [20] T. Norsen, Found. Phys. **39**, 273 (2009).
- [21] N. Gisin, Phys. Lett. A **154**, 201 (1991).
- [22] S. Popescu and D. Rohrlich, Phys. Lett. A **166**, 293 (1992).
- [23] R. Horodecki, P. Horodecki, M. Horodecki, Phys. Lett. A **200**, 223 (1996).
- [24] S. Yu, Q. Chen, C. Zhang, C. H. Lai and C. H. Oh, Phys. Rev. Lett. **109**, 120402 (2012).
- [25] J. Barrett, Phys. Rev. A **65**, 042302 (2002).
- [26] J. Degorre, S. Laplante and J. Roland, Phys. Rev. A **75**, 012309 (2007).
- [27] T. Maudlin, in *Biennial Meeting of the Philosophy of Science Association* (1992), pp. 404-417.
- [28] B. F. Toner and D. Bacon, Phys. Rev. Lett. **91**, 187904 (2003).
- [29] G. Murta, master's thesis, available at <http://www.mat.ufmg.br/~tcunha/DissGlaucia.pdf> (2012).
- [30] M. L. Almeida, S. Pironio, J. Barrett, G. Toth and A. Acin, Phys. Rev. Lett. **99**, 040403 (2007).
- [31] G. Toth and A. Acin, Phys. Rev. A **74**, 030306(R) (2006)
- [32] T. Vertesi and N. Brunner, arXiv:1405.4502 [quant-ph] (2014).
- [33] R. Augusiak, M. Demianowicz and A. Acin, arXiv:1405.7321 [quant-ph] (2014).